```
OM of: US-09-494-297-2 to: GenEmbl:*  out_format : pfs

Date: Jun 6, 2001 10:59 PM

About: Results were produced by the GenCore software, version 4.5,
       Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL-frame+_p2n.model -DEV-xlp
-Q-/cgn1_1/USPTO_spool/US09494297/runat_06062001_115734_12308/app_query.fasta_1.825
-DB-GenEmbl -QFMT-fastap -SUFFIX-pfs -GAPOP-12.000 -GAPEXT-4.000
-MINMATCH-0.100 -LOOPCL-0.000 -QGAPOP-4.500
-OGAPEXT-7.000 -XGAPOP-10.000 -XGAPEXT-0.500 -FGAPOP-6.000
-FGAPEXT-7.000 -YGAPOP-10.000 -YGAPEXT-0.500 -DELOP-6.000
-DELEXT-7.000 -START-1 -MATRIX-blosum62 -TRANS-human40.cd1
-LIST-45 -DOCALIGN-200 -THR_SCORE-pct -THR_MAX-100 -THR_MIN-0
-ALIGN-15 -MODE-LOCAL -OUTFMT-pfs -NORM-ext -MINLEN-0
-MAXLEN-200000000 -USER-US09494297_@CGN1_1_7604 -LONGLOG -NO_XLPXY -WAIT -THREADS-1
-ICPU-3 -LONGLOG -NO_XLPXY -WAIT -THREADS-1

Search information block:
Query: US-09-494-297-2
Query length: 757
Database: GenEmbl:*
Database sequences: 1283235
Database length: 1216004940
Search time (sec): 2987.740000
```

score_list:

Sequence	Strd	Orig	ZScore	EScore	Len	! Documentation
gb_ba3:SPU49397	-	1965.50	1843.52	1.8e-94	10826	! U49397 Streptococcus pyogenes
gb_ba3:STRPRTNF	+	633.50	597.92	4.3e-25	2845	! L10019 Streptococcus pyogenes
gb_ba1:AF009908	+	611.50	576.82	6.5e-24	2801	! AF009908 Streptococcus pyogene
gb_ba3:SPSFB	+	611.00	571.10	1.4e-23	2263	! X67947 S. pyogenes Sfb gene for
gb_ba3:GGU31115	+	523.00	496.38	2.0e-19	2066	! U31115 Group G streptococcus f
gb_ba1:AF009914	+	272.50	275.93	3.7e-07	321	! AF009914 Streptococcus pyogenes
gb_ba1:AF009913	+	262.50	266.71	1.2e-06	313	! AF009913 Streptococcus pyogenes
gb_ba1:AF009910	+	241.50	246.54	1.6e-05	326	! AF009910 Streptococcus pyogenes
gb_ba1:AF009920	+	222.50	228.05	0.0002	348	! AF009920 Streptococcus pyogenes
gb_ba1:AF009917	+	211.50	219.59	0.0005	278	! AF009917 Streptococcus pyogenes
gb_ba1:AF009911	+	211.50	218.66	0.0006	310	! AF009911 Streptococcus pyogenes
gb_ba1:AF009918	+	211.50	217.74	0.0006	309	! AF009918 Streptococcus pyogenes
gb_ba1:AF009915	+	209.50	216.80	0.0007	309	! AF009915 Streptococcus pyogenes
gb_ba1:AF009912	+	205.50	213.81	0.0011	282	! AF009912 Streptococcus pyogenes
gb_ba1:AF009919	+	204.50	212.11	0.0013	308	! AF009919 Streptococcus pyogenes
gb_ba1:AF009909	+	202.50	211.01	0.0015	281	! AF009909 Streptococcus pyogenes
gb_ba1:AF009916	+	185.00	195.32	0.0115	255	! AF009916 Streptococcus pyogenes
gb_ba2:SFPNGGEN	+	180.00	172.44	0.2170	2127	! X99995 S. equi fnz gene 2 from
gb_bal:AR083961	+	178.00	168.52	0.3584	2694	! AR083961 Sequence 2 from patent
gb_bal:AR083960	+	178.00	159.58	1.13	7650	! AR083960 Sequence 1 from patent
gb_bal:AF043131	+	178.00	158.24	1.34	8946	! AF043131 Moraxella catarrhalis
gb_pat2:AX067453	+	177.00	142.62	9.93	49617	! AX067453 Sequence 28 from Pat
gb_bal:AF043133	+	176.00	166.56	0.4611	2718	! AF043133 Moraxella catarrhalis
gb_ba3:SPU31980	+	164.00	152.94	2.64	3552	! U31980 Streptococcus pyogenes
gb_ba2:AP000992	-	162.50	113.00	443.49	318850	! AP000992 Thermoplasma volcan
gb_pr6:AP000503	-	162.00	112.56	469.73	100000	! AP000503 Homo sapiens genomi
gb_ba2:CDTOXBA	+	151.00	134.67	27.57	7164	! Z23277 C.difficile gene for to
gb_pr1:CDTOXBA	+	151.00	106.46	1.0e+03	192929	! AC005505 Plasmodium falcipar
gb_htg1:AC005139	-	151.00	104.03	1.4e+03	256172	! AC005139 Plasmodium falcipar
gb_htg1:AL139040	+	150.00	139.78	14.30	3531	! AR071281 Sequence 1 from paten
gb_bal:AR071281	+	150.00	139.39	15.05	3698	! AF071083 Streptococcus pyogene
gb_bal:AF071083	+	149.00	141.14	12.02	2700	! M58828 Human MHC class III HSE
gb_pr9:HUMMHHSP	+	149.00	137.88	18.25	3948	! AJ005646 Staphylococcus aureus
gb_ba3:SAU5646	+	149.00	136.55	21.65	4612	! AR067712 Sequence 8 from paten
gb_pat1:AR067712	+	148.00	137.21	19.90	3827	! M81736 Staphylococcus aureus d
gb_ba3:STACNA	+	148.00				
```
(continued)
```
gb_pat1:AR067706	+	148.00	137.21	19.90	3827	! AR067706 Sequence 1 from pa
gb_bal:AE002147	+	147.00	119.50	192.89	27093	! AE002147 Ureaplasma urealy
gb_pat1:A18434	+	146.00	136.16	22.75	3468	! A18434 Hybrid DNA molecule
gb_pat1:A18436	+	146.00	133.72	31.12	4612	! A18436 assembled sequence f
```

seq_name: gb_ba3:SPU49397

```
seq_documentation_block:
LOCUS       SPU49397    10826 bp    DNA             BCT       19-DEC-1998
DEFINITION  Streptococcus pyogenes MsmR (msmR) gene, partial cds; LepA (lepA),
            Cpa (cpa), and Nra (nra) genes, complete cds; SsbA (ssbA) gene,
            partial cds; and unknown genes.
ACCESSION   U49397
VERSION     U49397.1  GI:4028947
KEYWORDS    .
SOURCE      Streptococcus pyogenes.
  ORGANISM  Streptococcus pyogenes
            Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
            Streptococcus.
REFERENCE   1  (bases 1 to 10826)
  AUTHORS   Podbielski,A., Woischnik,M., Leonard,B.A.B. and Schmidt,K.H.
  TITLE     Characterization of nra, a global negative regulator gene in group
            A streptococci
  JOURNAL   Mol. Microbiol. 31 (1999) In press
REFERENCE   2  (bases 1 to 10826)
  AUTHORS   Podbielski,A. and Woischnik,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-FEB-1996) RWTH Aachen, Institute of Medical
            Microbiology, Pauwelsstrasse 30, Aachen, Nrw, Federal Republic of
            Germany, 52074
FEATURES             Location/Qualifiers
     source          1..10826
                     /organism="Streptococcus pyogenes"
                     /strain="CS101"
                     /db_xref="taxon:1314"
     gene            <1..630
                     /gene="msmR"
     CDS             <1..630
                     /gene="msmR"
                     /function="multiple sugar metabolism regulator"
                     /codon_start=1
                     /transl_table=11
                     /product="MsmR"
                     /protein_id="AAC97150.1"
                     /db_xref="GI:4033713"
                     /translation="DLVKLGNPQLLKQEINRIPLSSITSSSISALRAEKNLIYIYLTR
                     LEEFSFVENTDVAKHYSLVKYAMALNEBASDLLKVLRIRCAAIIFHFSESLTNKSISDK
                     ROMYNSVLHYVDSHLYSKLKVSDIAKRLYVSESHLRSVFKKYSNVSLQHYILSTKIKE
                     AQLLLKRGIPVGEVAKSLYFYDTTHFHFKIFKKYTGISSKDYLAKYRDNI"
     repeat_region   697..730
                     /note="putative transcription terminator"
                     /rpt_type=inverted
     CDS             complement(790..1377)
                     /note="ORF 2"
                     /codon_start=1
                     /transl_table=11
                     /product="unknown"
                     /protein_id="AAC97151.1"
                     /db_xref="GI:4033714"
                     /translation="MRKYWKMLFSVVNMLTMLAFNQTVLAKDSTVQTSISVENVLERA
                     GDSTPFSIALESIDAMKTEEETIIAGSGKASFSPLNFTTVGQYTRYVQKPSONKDIQ
                     ADTTVFDVLVYVTYDEDGTLVAKVISRRAGDEEKSAITFKPRLVKFIPPRQPDIPKT
                     PLPLAGEVKSLIGILSIVLLGLLVLLVYKKLKSRL"
     CDS             complement(1394..3169)
                     /note="similar to electron transfer flavoprotein 1 of
                     Methylophilus electron; ORF 3"
                     /codon_start=1
                     /transl_table=11
                     /product="unknown"
                     /protein_id="AAC97152.1"
                     /db_xref="GI:4033715"
                     /translation="MKKNLLLATAILATALGMASMSONIKAETAGVIDGSILVVKKT
                     FPSYTDDNVLMPKADYSFKVEADDNAKGKTKDGLDIKPGVIDGLENTKTIRYSNSDKI"
```

```
     gene            complement(3191..3712)
                     /gene="lepA"
     CDS             complement(3191..3712)
                     /gene="lepA"
                     /note="ORF 4; similar to Staphylococcus aureus signal
                     peptidase 1"
                     /codon_start=1
                     /transl_table=11
                     /product="lepA"
                     /protein_id="AAC97149.1"
                     /db_xref="GI:4033712"
                     /translation="MTNYLNRLNENSLFKAFIRLVLKISIIGFLGYILFQYVFGVMII
                     NTNDMSPALSAGDGVLYIYRLADRSHINDVVYEVDNTLKVGRIAAQAGDEVNFTQEGG
                     LLINGHPPEKVPYLYTYPHSSGPNPFYKVPGTYFIFLNDYREERLDSRYGALPINOI
                     KGKISTLLRVRGI"
     gene            complement(3705..5933)
                     /gene="cpa"
     CDS             complement(3705..6180)
                     /gene="cpa"
                     /note="collagen-binding protein of group A Streptococci"
                     /codon_start=1
                     /transl_table=11
                     /product="cpa"
                     /protein_id="AAC97148.1"
                     /db_xref="GI:4033711"
                     /translation="MQKRDKTNYGSANNKRRQTTIGLLKVPLTFVALIGIVGFSIRAF
                     GAEEQSVPNRQSSIQDYPWYGYDSYPKGYPDYSPLKTYHHNLKVNLEGSKDYQAYCFNL
                     TKHFPSKSDSVRSOWYKRLEGTNENFIKLADKPRIEDGQLOQNIRLIYNGYPNNNG
                     IMKGIDPLNAILVTQNAIWTDSAQINPDESFKTEARSNGTINDQQLGIMRKALKELID
                     PNLGSKYSNKTPSGIRLNVFESHDKRPFQNLLSAEYVPDTPPKPGEEPPAKTEKTSVII
                     RKYAEGDSKLLEGATLKLSQIEGSGEQEKDEQOSNSLGGETVELPNGTYTLTETSSPDGY
                     KIAEPIKFRVENKKVFIVQKDGSQVENPNKEVAEPYSVEAVNDFMDEEVLSGFTPYGK
                     FYAINKDKSSQVVCFNADLHSPPDSYDSGGETINPDTSTMKEVKTYHTAGSDLFKYA
                     LRPRDYNPEDFELRKHVKIERGYKKRQDSNGLTEQFRRAATQLAIYYFDSADLKTL
                     KTYNNGKGYHGFESMDEKTLAVTKELITYAQNGSAPOLTNLDFFVPNNSKPQSLIGTE
                     CHPDLDVDVIRMEDKKQEVIPVTHSLTVKKTVVGELGDKKGFQFELELKDKTGQPIV
                     NTLKTNNQDLVAKDGKYSFNLKHGDYIRIEGLPTGYSYTLKEAEAKDYIVTAVDNKVSQ
                     EAQSVGKDIEDKKVTFENRDLVPPTGLITDGAIIWLILLVPLGLLVWLFGRKGLK
                     ND"
     -10_signal      complement(5972..5977)
                     /gene="cpa"
                     /note="promoter 1"
     -35_signal      complement(5997..6102)
                     /gene="cpa"
                     /note="promoter 1"
     -10_signal      complement(6138..6148)
                     /gene="cpa"
                     /note="promoter 1"
     -35_signal      complement(6161..6166)
                     /gene="cpa"
                     /note="extended promoter 2"
     -10_signal      6231..6241
                     /gene="cpa"
                     /note="promoter 2"
     -10_signal      6231..6241
                     /gene="nra"
                     /note="extended promoter 2"
     gene            6231..7898
                     /gene="nra"
     -35_signal      6297..6302
                     /gene="nra"
                     /note="promoter 1"
     -10_signal      6315..6325
                     /gene="nra"
                     /note="promoter 1"
     CDS             6363..7898
                     /gene="nra"
```

```
                     /note="negative regulator of group A streptococci"
                     /codon_start=1
                     /transl_table=11
                     /product="Nra"
                     /protein_id="AAC97147.1"
                     /db_xref="GI:4028948"
                     /translation="MPVVRKVKDSFLVETYLEQSIRDKSELVLLLFKSPTIIFSHVAK
                     QTGLTAVOLKYCKELDDFFGGNNLDITIKKGKIICCFVKPVKEFYLHQLYDFSTIKL
                     SVAMGSLFDTRNIHDFDYSLIYGHHMAGNAMFGEIPKFLKNFFNKIHNKAIIET
                     IVFFIKNGTSSOPLIFKFSKKYKFLSSSSAYRIRESLIKLLREGLRVSKNTIVGEERI
                     RYLIAMIYSKFGIVIYPLDHLDNQIIYRFLSQSATNLRISPWLEEPSFYNMLLALSW
                     KRHQFAVSIPQTRIFRQLKKLFIYDCLIRSSRQVIENAFSLIFSQGDLDYLFLIYITT
                     NNSFASLQWTPOHIETCCHIFEKNDFRLLLEPILKRLPQINHSKQDLIKALMYFKRS
                     FLPNLQHFVIEIPSFSLPTYTGNSNLYKALRNIVNQWLAQLPGKRHLNEKHLQLFCSH
                     IEQILKNKOPALIVVLISSNFINAKLLfTDTIPRYFSDKGIHFYSFYLLRDIIYQIPSL
                     KPDLVITHSRLIPFVKNDLVKGVTVAEFSFDKRPDYSIASIONLIYQLKDKKYQDFLNE
                     QlQ"
     -10_signal      8006..8016
                     /note="ORF 5; extended -10 signal"
     CDS             8024..8857
                     /note="ORF 5"
                     /codon_start=1
                     /transl_table=11
                     /product="unknown"
                     /protein_id="AAC97154.1"
                     /db_xref="GI:4033717"
                     /translation="MDKIIKSIAQSGPRAYAVLDSTETVALAQEKHNTLSSSTVALGR
                     TLIANQILAAANQKGDSKITVKVIGDSSFGHIISVADPKGHVKGYIQNTGVDIKKTERG
                     EVLVGPEMGNGHFVTIIDYGTGNPYTSTTPLLITGIEGIEDFAYYLTESEQTPSAIGLNV
                     LLDENDKVKVAGGFMVQVLPEASEEEIARYEKRLQEMPAISHLLASKNHVDALLEAIY
                     GDEPYKRLSEEPLSFQCDCSRERFEAALMTLPKADLQAMIDEDKGAEIVCQFCGTKLP
                     I"
     gene            8884..9861
                     /note="ORF 6; similar to Azospirillum brasilense nifR3
                     gene"
     CDS             9881..10522
                     /codon_start=1
                     /transl_table=11
                     /product="unknown"
                     /protein_id="AAC97155.1"
                     /db_xref="GI:4033718"
                     /translation="MIKLNSSFRIGDVEIPHRIVLAPMAGVTNSAFRTIAKEFGAGLV
                     VMEMISEKGLLYNNEKTLHMLHIDENHPMSIOLFGGDAEGIKRAADFIQTNKADIV
                     DINMGCPVNKVVKNEAGAKWLRDPDKIYIHVKEVTSVLDIPLTVKMRTGWADSSLAVE
                     NALAAESAGVSALAMHGRTREQWYTGCDHITEARVASKAITHIPFIGNGDVRSVODAK
                     FMIEEIGVDAVMIGRAANNPYLFTQINHFFEEGQELPDLPFAKKLDIAKDHLKRLIN
                     LKGETIAVREFRGLAPHYLRGTAGAAKVRGAVSRAEFLAEVEAIFETVR"
     repeat_region   10598..10681
                     /note="putative transcription terminator"
                     /rpt_type=inverted
     gene            complement(10782..>10826)
                     /gene="ssbA"
     CDS             complement(10782..>10826)
```

```
alignment_scores:
     Quality:          1965.50
     Ratio:            3.412
Percent Similarity:    78.155          Percent Identity:    54.410
alignment_block:
     US-09-494-297-2 x SPU49397/rev
     US-09-494-297-2 x SPU49397/rev
     Length:    737
     Gaps:      13

Align seq 1/1  to reverse of: SPU49397  from: 1  to: 10826
```

```
  20 SerLysAsnSerLysArg......PheThrValThrLeuValGlyValPh  34
     |||::|||::::|||||||
5903 AGGGCTAACAACAAACGACGACAAACGACGGATACGGATTACTGAAAGTATT 5854

  34 eLeuMetIlePheAlaLeuValThrSerMetValGlyAlaLysThrValP  51
     |||||                              ::::
5853 TTTGACGTTTGTTAGCTCTGATAGGAANAGTAGGGTTTCTATCAGACGCGT 5804

  51 heGlyLeuValGluSerSerThrProAsnAlaIleAspProSerSer     67
     |||||                             ::::|||
5803 TCGGAGGCTGAAGAACACAATCAGTAGCACCAAAT......AGACAAAGCTCAATT 5760

  68 SerGluTyrArgTrpTyrGlyTyrGluSerTyrValArgGlyHisProTy  84
     |::::|||||||||
5759 CAAGATTATCCGTGGTATGGGCTATGATTCTTATCCTAAAGGCTACCCAGA 5710

  84 rTyrLysGlnPheArgValAlaIleHisAspLeuArgValAsnLeuGluGlys 101
     |||::::
5709 CTATAGCTCCGTTAAAGACTTACCATAATTTAAAAGTAAATTTAGAGGGAA 5660

 101 erArgSerTyrGlnValTyrCysSerPheAsnLeuLysLysAlaPheProLeu 117
     |::::|||||::::|||||:::                   ||||||
5659 GTAAGGATTATCAAGACATACTGCTTTAATTTAACAAAACATTTTCCATCC 5610

 118 GlySerAspSerSerValLysLysTrpTyrLysLysHisAspGlyIleSe  134
     |||||||                          :::|||||
5609 AAGTCAGATAGTTGTTAGATCACAATGGTATAAAAAAACTTGAAGGAACTAA 5560

 134 rThrLysPheGluAspPhrValAlaMetSerProArgIleThrGlyAspGlu  151
     :::||||                         ::|||
5559 TGAAAACTTTATCAAGTTAGCAGATAAAACCAAGAATAGAACGGGACGCAGT 5510

 151 euAsnGlnIleLysSerLeuArgAlaValMetTyrArgAsnGlyHisProGlnAsnAla 167
     |||||||||:::::                     |||||
5509 TACAACAAAATATATTGAGGAATTCCTCTAATAATGGATAATCCTAATAAATCGT 5460

 168 AsnGlyIleLeuMetGluGlyLeuGluProLeuAsnAlaIleLeuArgValThrGl  184
     ||||||||||                     ::|||
5459 AATGGGGATAATGGAAAGGATAGNTCCCTCTAAAACGCTAATTTTACTGACTCA 5410

 184 nGluAlaValTrpThrTyrTyrSerArgAspAsnAlaProIleSerArgAsnProAspG 201
     :::|||:::|||                   |||||||
5409 AAAATGCTATTTGG....TATACTGATTCAGCTCAAAATT...AAATCCGGATG 5366

 201 luSerPheLysArgGluSerGluAsnLeuValSerThrSerGlnLeu 217
     ||:::||||||                   :::||||||
5365 AAAGTTTTTAAAACAGAGAGCCTCGAAAGTAATGGTATTAATGACCAGCAGTTA 5316

 218 SerLeuMetArgGlnIleAlaLeuLeuLysGlnLeuLeuIleAspProAsnLeuAlaTh 234
     :::|||||||||                   ::||||
5315 GGCTTAATGCGACAAAAAGCTTTAAAAGAACTAATTGATCCAAACTTAGGGTC 5266

 234 rLysMetProLysGlnValProAspAspSerAspPheGlnLeuSerIlePheGlus 251
     |||                            ::||||
5265 AAAAATATTCCGAATAAAAACTCCATCAGGTTATCGGTTAAAATGTATTGAAT 5216

 251 erGluAspLysAspLysGlyAspLysTyrAsnLysGlyLeuTyrGlnLeuAsnLeuSer 267
     |::::|||||                      ::||||||
5215 CTCAGTGAT... AAGCCTTTCCAAAATCTTTTGAGT 5184

 268 GlyGlyLeuGlyValProThrPheLysProProProGlyAspProMetPr 284
     |||||                          ||
5183 GCTGAGTATGGTTCCGGAANCTCCCCCAAAACCAGGAGAGAG......CC 5140

 284 oProAsnGlnProGlnThrThrSerValLeuIleArgLysTyrTyrAlaIleG 301
     ||||:::                          |||||||
5139 TCCGGCTAAAACTGAAAAAACATCAGTCATTATCAGAAAAAATATGCGGAAG 5090

 301 lyAspTyrSerLysLeuLeuGluGlyLeuAlaThrLeuGlnLeuThrGlyAsp 317
     |||||||||                      ||||||||||
5089 GTGAC...TCTAAAACTTCTAGAGGAGGAGCAAACCTTTAAAGCCTTCTCAAATT 5043

 318 AsnValAsnSerPheGlnAlaArgValPheSerSerAsnAspIleGlyGl 334
     ||||||||||||::::||||
```

```
5042 GAAGGAAGTGGTTTTCAGAGAAAAAGGACTTTCAAAGTAATAGTTTAGGAGA 4993
     :::|||||||                   |||||::::|||||||

 334 uArgIleGluLeuSerAspGlyThrTyrThrLeuThrGluLeuAsnSerP 351
     |||::::
4992 AACTGTCGAATTACCAAAATGGGACTTTATTACCCTTAACAGAACATCTC 4943

 351 roAlaGlyTyrSerIleAlaGluIleProIleThrPheLysValGluAlaGly 367
     |::::                          ||||::::|||||||
4942 CAGATGGATATAAAATTGCGGAGCCGGATTAAGTTGAGCTAGAGAATAAA 4893

 368 LysValTyrThrIle...IleAspGlyLysGlnIleGluLeuAsnProAsnLy 383
     |||                           :::|||||||||
4892 AAAGTATTTATCCGTCCAAAAAGATGGTTCTCAAGTGGAAAATCCAAACAA 4843

 383 sGluIleValGluProThrTyrSerValGluAlaIleTyrAsnAspPheGluGluP 400
     |||||||                      ||||||||||
4842 AGAAGTAGAGCAGAGGCCAATACTCAGTTGGAAGGCGTAATAAATGACTTTGAGGATG 4793

 400 heSerValLeuThr...ThrGlnAsnTyrAlaLysPheArgTyrAlaLys 415
     :::|||||                      ||||::::|||||||
4792 AAGAAGTACTCTCGGGTTTTTACTCCATACGGAAAAATTCTATTACGCTACA 4743

 416 AsnLysAsnGlySerGlnValValThrCysPheAsnAlaAspLeuLeu 432
     ||::::                        ||||||||||
4742 AATAAGGATAAAAGTTCACAAGTTGTCACTTGCTTCAATGCTGATTTACA 4693

 432 LysTyrThrValLysProArgAspThrAsnProThrPheLeuLysHi 449
     |||||||                      ::||||||
4692 AAATATGCCCTAAGACCGAGAGATACAAACCCAGAGACTTTCTTTAAAGCA 4643

 449 hrThr...GlyGluValLysTyrThrHisIleLeuAlaGlyArgAspLeuPhe 464
     :::|||                         ||||||||||
4642 GTACGGATGAAAGAAGTCAAGTACACACATATACGGCAGGTAGTGACTTGTTT 4593

 465 LysTyrThrValLysProArgAspThrAsnProThrPheLeuLysHi 481
     |||||||                      ::|||||||
4592 AAAATATGCCCTAAGACCGAGAGATACAAACCCAGAGACTTTCTTAAAGCA 4543

 481 sIleLysLysValIleGluGluLysGlyGlyTyrArgGluLysGlyGlnIleAlaIleGlu 498
     :::|||                        ||||||||||
4542 CATTAAAAAAGTAATTGAAGAAAAGGGCTACAGAAAAAAGGGTGATAGC... 4497

 498 luTyrSerGlyLeuThrGlyIleGlnLeuGluArgAlaLeuAlaThrGluLeuAla 514
     |||:::                        ||||||||||
4496 ..TATAATGGATTAACAGGAAACACAGTTTCGCCGCGGCTACTCAGCTTGCT 4449

 515 IleTyrTyrPheThrAspSerArgAlaGlyLeuLeuAspLysLysSerLeu...... 529
     |||||||                       |||
4448 ATCTATTATTTTACAGACAGTGGCTGACTTAAAAACTTAAAAACTTATA 4399

 530 ......LysAspTyrHisGlyPheGlyLysAspMetLysAsnAspSerThrLeuA 544
     :::|||                        ||||||||||
4398 CAATGGGAAAGGTTACCATGGATTTGAANTCTTATGGATGAAAAAACCCTAG 4349

 544 laValAlaLysIleLeuLeuValGluTyrAlaGlnAspSerAsnProGln 560
     |||                           ::||||||
4348 CTGTCACAAAAGAATTAATTACTTTATGCTCAAAATGGCAGTGCCCCTCA 4299

 561 LeuThrAspLeuAspPheHeIleProAsnAsnAsnLysTyrGlnSerIle 577
     |||||||                      ||||||||||
4298 CTAACAAAACTTGANTTCTTCCGTACCTAATAAATAGCCAAAGACCAANCTCT 4249

 577 uIleGlyThrGlnIleTrpHisProGluGlnLysLysTyrArgMetG 594
     ||||||                        |||||
4248 TATTGGGACAGAAATCCATCCAGATGCATTTTGGGTTGACGTATTCGTATGG 4199

 594 luAspLysLys...GluValIleProValAThrHisHisAsnLeuThrLeuArg 609
     |||||||                      ||||||||||
4198 AAGATAAAAAAGCAAGAAGTTATTCCAGTAACTCACAGTTTGACAGTGAAA 4149

 610 LysThrValThrGlyLeuValAlaGlyAspArgThrLysAspPheHisPheG1 626
     |||||||||||::::||||
```

seq_name: gb_ba3:STRPRTNF

```
LOCUS       STRPRTNF       2845 bp     DNA            BCT       24-NOV-1993
DEFINITION  Streptococcus pyogenes fibronectin-binding protein (prtF) gene,
            complete cds.
ACCESSION   L10919
VERSION     L10919.1  GI:425479
KEYWORDS    adhesin; fibronectin-binding protein; protein F; prtF gene.
SOURCE      Streptococcus pyogenes (strain JRS75) DNA.
ORGANISM    Streptococcus pyogenes
            Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
            Streptococcus.
REFERENCE   1  (bases 1 to 2845)
AUTHORS     Sela,S., Aviv,A., Tovi,A., Burstein,I., Caparon,M.G. and Hanski,E.
TITLE       Protein F: An adhesin of Streptococcus pyogenes binds fibronectin
            via two distinct domains
JOURNAL     Mol. Microbiol. 10, 1049-1055 (1993)
MEDLINE     95020565
FEATURES             Location/Qualifiers
     source          1..2845
                     /organism="Streptococcus pyogenes"
                     /strain="JRS75"
                     /db_xref="taxon:1314"
     gene            425..2404
                     /gene="prtF"
     CDS             425..2404
                     /gene="prtF"
                     /codon_start=1
                     /transl_table=11
                     /function="binds fibronectin"
                     /product="protein F"
                     /protein_id="AAA26964.1"
                     /db_xref="GI:425480"
                     /translation="MFVRGEKMNKIFLNKEASFLAHTKRKRRFAVTLVGVFFMLLAC
                     AGIGFGQVAYAADEKTVPSHSHSPNPEPPWYGDAYGKEYPGYNIWTRYHDLRVNLNG
                     SRSXQVYCFNIQSNYPSQKNSFIKNWFKKIEGNGKSFVDYAHTTRLGKEELEQRLLSL
                     LYNAYPNDANGYMKGLEHLNAITVTQYAVWHYSDNSQYQFFTLWESEAKEGKISRSQV
                     TLMREALKKLIIDPNLEATAVNKIPSGYRLNIFFSSENEAYQNLLSAEYVPDDPRKPGET
     repeat_unit     1327..1457
                     /gene="prtF"
                     /note="repeat domain 1"
     repeat_unit     1631..2170
                     /gene="prtF"
                     /note="repeat domain 2"
BASE COUNT      977 a    472 c    616 g    780 t
ORIGIN
```

```
SEHNPKTPELDGTPIPEDPKHPDDNLEPTLPPVMLDGEEVPSESLEPALPLLMPE
LDGQEVPEKPSIDLPIEVPRYEFNNKDQSPLAGESGETEVITEEVYGNQPVDIDKKL
PNETGFSGNMVETEDTKEPEVLMGGQSESVEFTKDTQTGMSGQTTPQVETEDTKEPEV
LMGGQSESVEFTKDTQTGMSGQTTPQIETEDTKEPEVLMGGQSESVEFTKDTQTGMSG
QTTPQVETEDTKEPEVLMGGQSESVEFTKDTQTGMSGQTTPQIETEDTKEPEVLMGGQ
SESVEFTKDTQTGMSGPSETATVVEDTRPKLVHFDNNEPKVEENREKPTKNITPILP"
```

alignment_block:
US-09-494-297-2:

alignment_scores:
       Quality: 633.50   Length: 779
       Ratio: 1.560   Gaps: 30
Percent Similarity: 52.118   Percent Identity: 26.059

Align seq 1/1 to: STRPRTNF   from: 1   to: 2845

```
20  SerLysAsnSerLysArgPheThrValThrLeuValGlyValPheLeuMe 36
    ::::|||::|||:::|||||||||||||||||||||||||||||||||
494 ACAAAAAGAAAAAAGGCGATTGCTGCTACTTTAGGTGGGGAGTCTTTTAT 543

36  tIlePheAlaLeuValThrSerMetValGlyAlaLaLysThrValPheGlyL 53
    :::::||||||||||||::|:|::::||||||||||||||:||||:||
544 GCTTTTGCCATGCGCGGCGGTGCTATCGGTTTTGGTCAACTAGCCTATCTG 593

53  euValGluSerThrProAsnAlaIleAsnProAspSerSerSerGlu 69
    :::|||||||||||||||||||:::|:||:|||:::||||||||:|||
594 CCGATGAGAAGACAGTGCCGAGTCATACTAGTAGTCCA......AATCCTGAG 637

70  TyrArgTrpTyrGlyLeuTyrGluSerTyrValAlaArgGlyHisProTyrTyrLy 86
    :::|||||||||||||||||||||||:::||||:||:|||||| ::|:
638 TTTCCTTGGTACGGTTATGATGGTTAAGGAATATCCTGGATATAA 687

86  sGlnPheArgValAlaHisAspLeuArgValAsnLeuGluGlySerArgS 103
    :::||||||||||||:|||||||::||:|||||||:|||||||||||
688 TATTTGGACTCGTTATCATGACTTGACAGTAAAATTAAATGCAGGGA 737

103 erTyrArgGlnValArgTyrCysPheAsnLeuGlyLysLysAlaAlaPheProLeuGlySer 119
    :::::|||::||:||::||||||||||::::||||||:::||::|||
738 GCTACCAAGTTATTGTTTTAATTATCAATCTAATTATCCTAGCCAAAA 787

120 AspSerSerValLysLysTrpPheTyrLysLysHisAspGlyTyrIleSerThrArg 136
    :::::|||||||||||:||||||||||||||:|||:|::||||||:||
788 AATAGTTTCATAAAAAATTGGTTTAAAAAAATTGAGGGAAATGGCAAATC 837

136 sPheGluAspTyrAlaMetSerProArgIleThrArgLeuGlyAspGluLeuAsnG 153
    :|||    |||||||||:::::||||:|||:::::|||||||||:|||
838 CTTTGTAGATTATGCTCACACTACTAAACTAGGTAAAGAAGAGTTGGAGC 887

153 lnLysLeuArgAlaValMetThrTyrAsnGlyHisProGlnAsnAlaAsnGly 169
    |:::||:::|:|:::|||||||||||||:||||||:|::||||||||:||
888 AGAGGCTCTTAAGTTACTGAGTTTACTGTATAATGGTCATCCACAGAATGCT 937

170 IleMetGluGlyLeuGluProLeuAsnAlaIleArgValThrGlnGluAl 186
    :::||:::|||||||||||||||:||:|||||||:|||||::||||
938 TATATGAAAGGACTGGAACCATTTAAAATCCTATTACTGTAACACAATATGC 987

186 aValTrpTyrTyrSerAspAsnAlaProIleSerArgAsnProAspGluSerP 203
    ||||||:::|||||||:|::|||||||:::||||||:::||||||:::
988 AGTTTGGCATTACTCAGATAATTCACAATATCAATTA... GAGACATTAT 1034

203 heLysArgGluSerGluAsnLeuValSerThrSerGlnLeuSerLeu 219
    :::|||:::||:::::::||:::::||||:::::||||||||||:|||
```

seq_documentation_block:
seq_name: gb_ba3:STRPRTNF

```
4148 AAAACAGTAGTCGGTGAGTTGGGAGATAAAAACTAAAGGCTTCCAATTTGA 4099
     ::|||||||||||::::|||||||||||||||||||::|||||||||||
626  uIleGluLeuLysAsnAsnLysGlnGlyLeuLeuSerGlnThrValLysT 643

4098 ACTTGAGTTGAAAGATAAAAACTGGAACAGCCTATTGTTAACACTCTAAAA 4049
     :::|||||||||||||||||||||||::|||||||||||||||||||
643  hrAspLysThrAsnLeuGluPheLysAspGlyLysAlaThrIleAsnLeu 659

4048 CTAATAATCAAGATTTAGTAGGTAAAGATGGGAAATATTCATTTAATCTA 3999
     ||||||||||||||::::::|||||||||||||||||:||||||||
659  LysHisGlyGluSerLeuThrLeuGlnGlyLeuProGlyThryrSerTy 676

3998 AGCACAGGTGACACCATAAGAATAGAAAGGATTATACCGACGGGATATCTTA 3949
     ||||||||||||||::::|||||||||||||||::||||||||||
676  rLeuValGlyGluThrAspSerGluGlyTyrIleArgLysValAlaAsnSerG 693

3948 TACTCTGAAACAGGCTGAAAGCTAAGATAATATAATAGTACCGTTGAATAACA 3899
     ||||||||||||::::|||||||||:::||||||||:|||||||||||
693  lnGluValAlaAasnAlaThrValSerLysThrGlyIleThrSerAspGlu 709

3898 AAGTTAGTTCAAGAAGCCCAGTCAGTAGGTAAGATATAACAGAAGACAA 3849
     ::|||||||||||||||||:||||||||||:::||||||||||||
709  ThrLeuAlaPheGluLeuAsnAsnLysGluProValValProThrGlyValAlaS 726

3848 AAAGTCACTTTTGAAAAACCGAAAAAGATCTTGTCCCACCAACTGGTTGAC 3799
     :::||||||:::||||||||||||||||||||::::::||||||:||
726  pGlnLysIleAsnGlyTyrLeuAlaLeuLeuIleValIleAlaGlyIleLeuSerL 743

3798 AACAGAGGGGCTACTCATCTTGGTGTTGTTAATTACTTGTTCCACTTGGT 3749
     ||:::| ||||||||||:|||:|||::||||||||||||::|||:||
743  euGlyIleTrp 746

3748 TATTGGTTTGG 3738
     ||:::|||
```

```
1035 GGGAGTCTGAGGCGGAAAGAAGTTAAGAATTAGTAGGTCACAAGTTACTTTA 1084
 220 MetArgGlnAlaLeuLeuLysGlnLeuIleAspProAsnLeuAlaThrLysMe 236
     ||||||::|||||||:|||||||||||||||||||||||||||||||
1085 ATGCGTCAGGCATTGAAAAAAAACTAATTGATCCCAAATTAGAAGCTACTGC 1134

 236 trpProLysGlnValProAspAspPheGlnLeuSerIlePheGluSerGluA 253
     |||:|||||||||:::|::|||||||||:||||||
1135 AGTTAATAAATCCATCAGGCTATCGTTTAAATATCTTTGAGTCTGAA.    1183

 253 spLysGlyAspLysTyrAsnLysGlyTyrGlnAsnLeuLeuSerGlyGly 269
     |||||||||||||:|||||||||||||:::
1184 ...........AATGAAGCTTACCAAAAATCTTTTAAGTGCTGAA      1216

 270 LeuValProThrLysProGlyThrProGlyIleAspProProMetProProAs 286
     |||||||      |||||||:
1217 TAAGTACCTTGATGATCCCCCTAAACCTGGTGAG.....ACGTCAGAACA  1260

 286 nGlnProGlnThrThrSerValLeuIleArgLysTyrAlaIleGlyAspT 303
1261 TAATCCTAAAACTCCCGAG.............................    1279

 303 yrSerLysLeuLeuGluGlyAlaThrLeu..........GlnLeuThr  315
     |||:|||||:||                    ::
1280 ...........TTGGATGGCACTCCAATTCCCGAGGACCCAAAACATCCA  1318

 316 GlyAspAsnValAsnSerPheGlnAlaArgValPheSerSerAsnAspIl 332
     ||||||:::|||:
1319 GATGATAATTAGAACCTACATTGCCCGGCCAGTGATGCTAGAT......  1360

 332 eGlyGlnArgIleGluLeuSerAspGlyThrTyrThrLeuThrGluLeuA 349
1361 .GGTGAACAAGTC............................CCAGAAGTTC 1382

 349 snSerProAlaGlyTyrSerIleAlaGluProIleThrPheLysValGlu 365
     |||:||                                ::
1383 CAAGCGGAGCTTAGAGCCTGCGCTTCCCCATTAATGCCAGAG........ 1426

 366 AlaGlyLysValTyrThrIleIleAspGlyLysGlnIle...GluAsnPr 381
     :||||||||||:::||:::|||      |||::|||
1427 ..........TTAGATGGTCAAGAAGTTCCCTGAAAAACC          1455

 381 oAsnLysGluIleValGluProTyrSerValGluAlaTyrArgAsnAspPheG 398
     ||:||            |||:||
1456 TACTATTGACTTA.......CCTATTGAAGTTCCTCGTTATGAG......  1492

 398 luGluPheSerValLeuthrThrGlnAsnTyrAlaLysPheTyrTyrTyrAla 414
     |||||                    |||
1492 .....AAGTGTTGATGGGAGGTCAAAGTGAGTCTGTTGAATTTACT.    1492

 415 LysAsnLysAsnGlySerSerGlnValValTyrCysPheAsnAlaAsple  431
     ||||||
1493 .................................TTTAACAATAAAGA      1506

 431 ulysSerProProAspSerGluAspGlyGlyLysThrMetThrProAspP 448
     ::|||||||            :::
1507 CCAATCACCTCTAGCGGGTGAG..........................    1528

 448 heThrThrGlyGluValLysTyr...ThrHisIleAlaGlyArgAspLeu 463
     ::|||||||||:|||          |||::::::
1529 .....TCTCGTGAGACGGAGTATATACCGAAGTCTATGGAAATCAA.    1570

 464 PheLysTyrThrValLysProArgAspThrProAspThrPheLeuLeuy 480
     ||||||                    |||
1571 ..........CAGAACCCTGTTGATATTGATTAAA.............    1594

 480 shIsIleLysLysValIleGlyValTyrArgGluLysGlyGlyGlnAlaI 497
     |||                              :::
1595 ..........AAACTTCCGAATGAAACAGGTTTT.....TCAGGAAATAGG 1631

 497 leGluTyrSerGlyLeuThrGluThrGlnLeuuArgAlaAlaThrGlnLeu 513
     |||      :::||                    |||
1632 TTGGACAGAGAAGATACGAAAGACCGAGAAGTGTTGATGGGAGGCCAA... 1678
```

```
 514 AlaIleTyrTyrPheThrAspSerAlaGluIleLeuAspLysLysLeuuly 530
     ::|||||:||:|||||||||                    :|||||
1679 ...........AGTGAGTCTGTTGAATTTACTAAAGACACTCAAAC      1713

 530 sAspThrHisGly.................................... 534
1714 AGGCATGAGTGGTCAAACAACTCCTCAGGTTGAGACGAGAAGATACGAAAG 1763

 535 ...........PheGlyAspMetAsnAspSerThrLeuAlaValAla  546
                |||
1764 AGCCAGAAGTGTTGATGGGAGGTCAAAGTGAGTC.............    1798

 547 LysIleLeuValGluTyrAlaGlnAsnSer.................  556
     |||||||:||||||:||::||||:
1799 .........GTTGAATTTACTAAAGATACTCAAACACAGGCATGAGTGGTCA 1839

 557 .AsnProGlnLeuThrAspLeuAspPhePhelleProAsnAsnAsnL  573
     |||||:|              |||:::
1840 AACAACTCCTCAGATTGAGACAGAGAAGACAGGACCAGA.........  1882

 573 ysTyrGlnSerLeuIleGlyThrGlnTrpHisProGluAspLeuValAsp 589
1882 .................................................. 1882

 590 IleIleArgMetGluAspLysLysGluvalIleLeuProValThrHisAsnLe 606
     :||||||||:||||                    |||
1883 ...GTTGAATGGGAGGCCAAAGTGAGTCTGTTGAATTTACTAAAGATACT 1921

 606 uThrLeuArgLysThrValThrGlyLeuValAlaGlyAspArgThrLysAspP 623
     |||||||||:::::|||||:::                |||
1922 ...AAAGACACTCAAACAGGCCATGAGTGGTCAAAACAACTCCT.....  1960

 623 heHisPheGluIleLeuGluLeuLysAsnAsnAsnLysGln...GluLeuLeu 637
     ::|||||||                              |||
1961 ....CAGGTTGAGACAGAGAAGATACGAAAGACCCAAGAGTGTTGATG    2002

 638 .........SerGlnThrValLysThrAspLysThrAsnGluGluPheLy 651
     |||                                  :::
2003 GGGAGGTCAAAGTGAGTCTGTTGAATTTACTAAAGATACTCAAACAGGCAT 2052

 651 sAspGlyLysAlaThrIleAsnLeuLysHisGlyValSerLeuThrLeuG 668
     ::|||||:||||                      |||||
2053 GAGTGGTCAAACAACTCCCTCAGATTGAGACAGAGAAGAGCCAG      2102

 668 lnGlyLeuProGluGlyTyrSerTyrLeuValLysGluThrAspSerGlu 684
     |||      |||
2103 AAGTCTGTTGATGGGAGGTCAAAGTGAGTCTGTTGAATTTACT.....  2143

 685 GlyTyrLysValLysValAsnSerGlnGluValAlaAsnAlaThrValSe 701
     :::::|||                              ::
2144 .........AAAGACACTCAAACAGGCATGAGTGGTTTCAG          2175

 701 rLysThrGlyIleThrSerAspGluThr...............LeuAlaP  713
     ||:||:::                    :::
2176 TGAAAACGGGCGACCTGTTGTTGAAGATACGGGTCCGAAGTTAGTGTTCCATT 2225

 713 heGluAsnAsnLys.................................  717
     ||:||:||||||:::
2226 TTGACAATAATGAGCCCAAAGTGGAGAGAAGAATCGGAGAAAGCCTACAAAA 2275

 718 .......GluProValValPro....ThrGlyValAspGlnLeuIleAsnGl 731
     |||                    ||||||
2276 AATATAACACCTATCCTTCCTGCAACAGGAGATATTGAGAATGTTTTGGC  2325

 731 yThrTyrLeuAlaLeuIleValAlaileAlaGlyIleSerLeu 743
     ::|||::|||      ::||::
2326 CTTTCTTGGAATCCTTATTTTGTCAGTACTTTCTATT            2362
```

```
DEFINITION   Streptococcus pyogenes fibronectin/fibrinogen binding protein F
             (prtF15) gene, complete cds.
ACCESSION    AF009908
VERSION      AF009908.1 GI:2271466
KEYWORDS     .
SOURCE       Streptococcus pyogenes.
  ORGANISM   Streptococcus pyogenes
             Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
             Streptococcus.
REFERENCE    1  (bases 1 to 2801)
  AUTHORS    Katerov,V., Andreev,A., Schalen,C. and Totolian,A.A.
  TITLE      Protein F, a fibronectin-binding protein of Streptococcus pyogenes,
             also binds human fibrinogen: isolation of the protein and mapping
             of the binding region
  JOURNAL    Microbiology 144 (Pt 1), 119-126 (1998)
  MEDLINE    98129085
REFERENCE    2  (bases 1 to 2801)
  AUTHORS    Katerov,V.E., Andreev,A.S., Schalen,C. and Totolian,A.A.
  TITLE      Direct Submission
  JOURNAL    Submitted (20-JUN-1997) Med. microbiol, Lund University, Solvegatan
             23, Lund 22362, Sweden
FEATURES             Location/Qualifiers
     source          1..2801
                     /organism="Streptococcus pyogenes"
                     /strain="EF1949"
                     /serotype="M15"
                     /db_xref="taxon:1314"
     gene            450..2507
                     /gene="prtF15"
     CDS             450..2507
                     /gene="prtF15"
                     /codon_start=1
                     /transl_table=11
                     /product="fibronectin/fibrinogen binding protein F"
                     /protein_id="AAC38155.1"
                     /db_xref="GI:2271467"
                     /translation="MFLRGEKMNNKIFLNKEASFLAHTRKRRFAVTLVGVFFMLFAC
                     AGSIGFGQVAYAAADEKTENRINSNPEFFWYGYDSYSGRLLRYINLKVNGSKEYQAYC
                     FNLKRFEPKKEESSSPNWVKKLDGSTEFKKYAENPRSGEELRRHILKVLYNGYPNS
                     NEIMKGIDPLNAIIVTQNAIWYYSDSAPINDINNFFTSEANDLNIPPQQLTLMREALR
                     KLISSDENLVKQVPSNFKLSIFESSDKSYQNLLSAEYVPDDPKRGDTSEHNKPTEL
                     DGTPIPEDPKRPDESSGRALPLMPELDGEEVPEVPSESLEPAL PPLVPELNDGEEVPE
                     VPSESLEPALPPLMPELDGQEVPEVPSESLEPALPPLMPELDGEINPVEVPSESLEPAL
                     PPLMPELDGQEVPEKPSIDLPIEVPRYEENNKDQSPLAGESGEVFTKDTQTGMSGQTPVETEE
                     DIDKKLPNETGFSGSGNMVETEETKEPGVLMGGQSESVEFTKDTQTGMSGQTPQVETEE
                     TKEPGVLMGGQSESVEFTKDTQTGMSGQTAPQVETEDTKEPEVLMGGQSESVEFTKDTQTGSSGFSETVSIVEDTRPKL
                     QAGMSGQTAPQVETEDTKEPEVLMGGQSESVEFTKDTQTGSSGFSETVSIVEDTRPKL
                     VFHFDNNEPKVEEHREKPTKNITILPAIGDIENVLAFLGILILSVLSIFSLLKNKQN
                     NKV"
BASE COUNT     954 a   489 c   597 g   761 t
ORIGIN
```

```
20  SerLysAsnSerLysArgPheThrValThrLeuValGlyValPheLeuMe  36
    ::||||:||::.:|||||||||||||||||||||||||||||:||
519 ACAAAAAAGAAAAAGGCGAATTTGCTGTCACTTTAGTGGGAGTCTTTTTTAT  568

36  tilePheAlaLeuValThrSerMetValGlyValAlaLysThrValPheGlyL  53
    :||:||||||||  |||:||:      :|:||::||||||||||:
569 GCTTTTTGCATCGCGGCGGTTCTATCGGGTTTCGGTCAAGTAGCCTATCCTG  618

53  euValGluSerSerThrProAsnAlaIleAsnProAspSerSerSerGlu  69
    |||:|:..:.:                   |||    ::|||   |||
     GAGAGCTTAGAACCTGCG  1478
```

```
619  CGGATGAGAAGACTGAGCGCCGAATCGTATTAAT......TCAAATCCTGAG  662
     |||||||||||||||||||||||||||||||||       |||
70   TyrArgTrpTyrGlyTyrGluSerTyrValArgGlyHisProTyrTyrLy  86
     ::||||:|||.||||||||||||||||||||:|||||||||||:||:      |||
663  TTTCCTTGGTATGGGTATGATTCGTATGCGGGAAGATTGTTAAGATAT..  710

86   sGlnPheArgValAlaHisAspLeuArgValAsnLeuGluGlySerArgS  103
     ::::                   .CACAATCTGAAAGTAAAT.....  GGAAGTAAGG  738
711  erTyrArgGlnValTyrCysPheAsnLeuLysLysAlaPheProLeuGlySer  119
     ::||||||:||:||||||||||||||||||||||||||||||||         :::
739  AGTATCAAGCGTATTGTTTTAAACCTGAAAAAGGTTTGAGCGAAAAAGA  788

120  AspSerSerValLysLysTrpTyrTyrLysLysHisAspGlyTyrIleSerThrLy  136
     ||||:  ||||||                   |||||||||||||||||||:||
789  GAAAGTAGCCAGTCCTAATTGGTATAAAAATTAGATAGGGAGTACAGAAAC  838

136  sPheGlyAspTyrAlaMetSerProArgIleThrGlyAspGluLeuAsnG  153
     |||||  |||||:  .||||||||||||||||||||||||||||:::
839  TTTTTAAAAAATACGCTGAAAAACCCTAGGTTTAGTGGAGAGAAATTAAGAC  888

153  lnLysLeuArgAlaLavalMetTyrAsnGlyHisProGlnAsnAlaAsnGly  169
     |:||||||||||||||||||||:||||:|||||::::
889  GACATAATTGAAGGTTCTCTACAATGGATAATCCT...AATAGTAATGAA  935

170  IleMetGluGlyLeuGluProLeuAsnAlaIleIleAlaArgValThrGlnGluAlaI  186
     ::||||:|||||||||||||||||||||||||||||||:|||
936  ATAATGGAAGGGGATAGACCCTCTCAACGGCTATCTTAGTAACACAAAATGC  985

186  aValTrpTyrTyrSerAspSerAlaProIleSerAspAsnProAspGluSerP  203
     |:|||||||||||||||||||||||||||::::
986  CAtATGGTATTACTCGGATAGTGCACCGATTAATGATTAATTAATAACTTCT  1035

203  heLysMetProLysGlnValProAlaPheGlnLeuSerIlePheGlu  219
     |:||||||||||||||||:.:|||:||:|||||||||:||        |
1036 TTACTAGTGAAGGCCAATGAATCTTAATAATACCCAGCCAGCTAACGTTA  1085

220  MetArgGlnIleAlaLeuLysGlnLeuIle......AspProAsnLeuAlaTh  234
     :|:||||||||||||||||||||||||||      |||  ||||||||:||::
1086 ATGCGTGAAGCCTGAGAAAACTAATTCCTCTGATGAAAATTGGT..  1133

234  rLysMetProLysGlnValProAlaPheGlnLeuSerIlePheGlu  251
     |:||||||||||||:.:|||:|||:|||:|||||:||||
1134 AACAAGTTCCATCGAACTTTAAACTTAGTATTTTCGAGT  1173

251  erGluAspLysGlyAspLysTyrAsnLysGlyTyrGlnAsnLeuLeuSer  267
     :::|||||||||||::|:||:||::|||||||||||||||::
1174 AAGTCCTACCAAAAATCTTTTAAGT  1205
     CTTCTGAT....                                       1205

268  GlyGlyLeuValProThrLysProGlnThrProGlyAsp............  280
     |:||||                   |||||||:||
     GCTCGAATATGTACCTGATGATCCTCCTAAACCTGGTGATACGTCAGAACA  1255
1206 GCTCGAATATGTACCTGATGATCCTCCTAAACCTGGTGATACGTCAGAACA  1255

280  ....................ProProMetProProAsnGlnProGln  289
                         |||  :::||||||||||||||||||||
1256 TAATCCTAAAACTCCCGAGTTGGATGGCACTCCAAATTCCGAGGACCCAA  1305

290  ThrThrSerValLeuIleArgLysTyrAlaIleGlyAspTyrSerLysLe  306
     :::                       |||    :::
1306 AACGTCCAGATGAGAGTTCAGGAACCTGGGCTTCCCCCATTGATGCCAGAG  1355

306  uLeuGluGlyAlaThrGluGlnLeuThrGlyAspAsnValAsnSerPheG  323
     |||                |||  ::: |||:::||||||:||||
1356 CTAGATGGTGAAGAAGTCCCAGAGTTCCAAGCCGAGAGCGTAGAACCTGC  1405

323  lnAlaArgValPheSerSerAsnAspIleGlyIleGluArgIleGluLeuSer  339
     |||  ::|:||||:|||||:||||  :|:|||||
1406 GCTTCCCCCATTGGTGCCAGAGCTTAGAGGTGTGAAGAAGTCCCAGAAGTTC  1455

339  GAGAGCTTAGAACCTGCG  1478
1456 CAAGC...........                                    1478
```

```
340  .......              AspGlyThrTyrThrLeuThrGlyLeu    348
                         |||||||
1479 CTTCCCCCATTAATGCCAGAGCCTAGATGGC .CAAGAAGTCCCAGAAGT   1525

348  uAsnSerProAlaGlyThrTyrSerIleAlaGluProIleThrPheLysValG  365
                         |||.::
1526 TCCAAGGCGAGCTTAGAACCTGCGCTTCCTCCATTAATGCCAGAGCTAG   1575

365  luAlaGlyLysValTyrThrIle.   ......                    372
                         :::
1576 ACGGTGAAGAAAATCCCAGAAGTTCCAGAGGAGCAGCTTAGAAACCTGCGCTT 1625

373  ..............IleAspGlyLysGlnIle...GluAsnProAs       382
     :||||||||||:::|||||||:    |||:||||||::
1626 CCTCCATTAATGCCAGAGTTAGATGGTAAGAAGTCCCTGAAAAAACCTAG  1675

382  nLysGluIleValGluProTyrSerValGluAlaIleThrTyrAsnAspPheGluG 399
     ||||.:|:|:||||| |||:::
1676 TATTGACTTA.......CCTATTGAAGTTCCTCGTTATGAG           1709

399  luPheSerValLeuThrThrGlnAsnTyrAlaLysPheTyrTyrAlaLys  415
1709 ...............                                     1709

416  AsnLysAsnGlySerSerGlnValValThrTyrCysPheAsnAlaAspLeuVal  432
     :|||||| |||:||||||:
1710 ......TTTAACAATAAAGACCA                             1726

432  sSerProGluAspSerGluAspGlyGlyLysThrMetThrProAspPheT  449
     :||||||| ::||||
1727 ATCACCTTCTAGCGGGTGAG.......                         1745

449  hrThrGlyGluValLysTyr...ThrHisIleAlaGlyArg.........  461
     :|||||||||::: |||    |||
1746 .TCTGGTGACGGATTATATTATCGAAGTCTATGAGAAATCAACAGAAC    1793

462  ..........AspLeuPheLysTyrThrValLysProArgAspThrAspProAs 476
     |||::| |||    |||
1794 CCTGTTGATATTGATAAAAAACTT......CCGAATGAAACAGGTTTTC    1837

476  pThrPheLeuLysHisIleLeuLysLysSerValIleGluLysGlyTyrArgGluL 493
     :::
1838 AGGAAAATAGGTTGAGACGAGAAGAACAGCCAGGAGTTGATGG         1887

493  ysGlyGln........AlaIleGluTyrSerGlyLeuThrGluThrGlnLeu  507
     ||||||| |||:::
1888 GAGGCCAAAGTGAGTCGTTGATTTTACGAAAGACACTCAACCAGGCATG   1937

508  ArgAlaAlaThrGlnLeuAlaIleTyrTyrPheThrAspSerAlaGlIle  524
     ::: |||
1938 AGCGGTCAAACA.......ACTCCTCAGGTTGAGAC               1966

524  uAspLysLysAspLeuLysAspTyrHisGlyPheGlyAspMetAsnAsps  541
     .:::|: |||
1967 AGAAGACGAAAGAGACCAGGAGTGTTG..ATGGGAGGTCAAAGTGAGT    2013

541  erThrLeuAlaValAlaLysIleLeuValGluTyrValAlaGlnAspSerAsn 557
     ||
2014 CT......GTTGAATTTACGAAAGACACT.                      2036

558  ProProGlnLeuThrAspLeuAspPhePheIleProAsnAsnAsnLysTy  574
2036 .............                                      2036

574  rGlnSerLeuIleGlyThrGlnTrpHisProGluAspLeuValAspIle  591
     ||||||: ::::||||||:||
2037 .CAAACAGGCAATGAGTGGTCAAACAGCTCCTCAG.......G        2070

591  leArgMetGluAspLysLysGluValIleProValThrHisAsnLeuThr  607
     ::: |||||| |||
2071 TTGAGACAGAAGACACGAAAGAG............CCAGAAGTG        2102
```

```
608  LeuArgLysThrValThrGlyLeuAlaGlyLysAspArgThrLysAspPheHi  624
                    |||     |||
2103 ..........TTGATGGGAGGCGAGGCCAAAGTGAGTCGTTGA          2131

624  spPheGluIleLeuLysAsnAsnLysGlnGluLeuLeuSerGlnThrV   641
     |||:||: :::::|||||: ::::||||||:
2132 ATTTAAG......AAAGACACTCAAGCACCAGGCCATGAGCGGTCAAACAG  2172

641  aLysThrAspLysThrArgAsnLeuGluPheLysAspAspGlyLysAlaThrIle 657
     ::::|:|| ::::|||
2173 CT.......CCTCAGGTTGAGACAGAAGACACGAAAGAGCCAGAA        2210

658  AsnLeuLysHisGlyGluSerLeuThrLeuGlnGlyLeuProGluGlyTy  674
     |||
2211 GTGTTGATGGGAGGCGAGGCCAAAGTGAGTCGTTGAATTTACTAAAGAT   2255

674  rSerTyrLeuValLysGluThrAspSerGluGlyTyrLysLysValValA  691
     ::||| |||:::||||||:
2256 ......ACTCAAACAGGCCTCGAGCGCGGTTTCAGTGAAACAGTGA       2292

691  snSerGlnLeuValAlaAsnAlaThrValSerLysThrGlyIleThrSer  707
     ::::::::||:||||||||:||||||||
2293 GCAATGCTGTTGAAGAAAACGCGGTCCGAAGTTA............      2321

708  AspGluThrLeuAlaPheGluAsnAsnLys.  .....              717
     |||:||||||||:
2322 ......GTGTTCCAATTTGACAATAAATGACCCCAAAGTGGAAGAGCATCG  2365

718  ............GluProValValPro...ThrGlyValA            726
     |||:::|||| |||||
2366 GGAAAAGCCTACAAAAAAATATAACACCTATCCTTCCTGCAACAGGAGATA 2415

726  spGlnLysIleAsnGlyTyrLeuAlaLeuLeuValIleLeuAlaGlyValIleSer 742
     ::::::::||:|||::|:|||:::::::||||||||:::|||
2416 TTGAGAATGTGTTTGGCCTTTCCTTGGAAATCCTTATTTTGTCAGTACTTTCT 2465

743  Leu   743
     :::
2466 ATT   2468
```

seq_name: gb_ba3:SPSFB

seq_documentation_block:

```
LOCUS       SPSFB                   2263 bp    DNA             BCT       13-JUL-1995
DEFINITION  S.pyogenes Sfb gene for fibronectin-binding protein (partial).
ACCESSION   X67947 S42389
VERSION     X67947.1  GI:511149
KEYWORDS    fibronectin-binding protein; sfb gene.
SOURCE      Streptococcus pyogenes
  ORGANISM  Streptococcus pyogenes
            Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
            Streptococcus.
REFERENCE   1  (bases 1 to 2263)
  AUTHORS   Talay,S.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JUL-1992) S.R. Talay, GBF, Mascheroder Weg 1, 3300
            Braunschweig, FRG
  REMARK    revised by [4]
REFERENCE   2  (bases 1 to 1030)
  AUTHORS   Talay,S.R., Valentin-Weigand,P., Jerlstrom,P.G., Timmis,K.N. and
            Chhatwal,G.S.
  TITLE     Fibronectin-binding protein of Streptococcus pyogenes: sequence of
            the binding domain involved in adherence of streptococci to
            epithelial cells
  JOURNAL   Infect. Immun. 60 (9), 3837-3844 (1992)
   MEDLINE  92363585
REFERENCE   3  (bases 1 to 2263)
  AUTHORS   Talay,S.R., Valentin-Weigand,P., Timmis,K.N. and Chhatwal,G.S.
  TITLE     Domain structure and conserved epitopes of Sfb protein, the
            fibronectin-binding adhesin of Streptococcus pyogenes
  JOURNAL   Mol. Microbiol. 13 (3), 531-539 (1994)
   MEDLINE  95089690
REFERENCE   4  (bases 1 to 2263)
```

```
AUTHORS      Talay,S.R.
TITLE        Direct Submission
JOURNAL      Submitted (13-JUL-1994) S.R. Talay, GBF, Mascheroder Weg 1, 3300
             Braunschweig, FRG
COMMENT      On Jul 17, 1994 this sequence version replaced gi:47433.
FEATURES             Location/Qualifiers
     source          1..2263
                     /organism="Streptococcus pyogenes"
                     /strain="DSM 2071"
                     /db_xref="taxon:1314"
     sig_peptide     124..270
                     /gene="Sfb"
     gene            124..2040
                     /gene="Sfb"
     CDS             124..2040
                     /gene="Sfb"
                     /codon_start=1
                     /transl_table=11
                     /product="fibronectin-binding protein"
                     /protein_id="CAA48133.1"
                     /db_xref="GI:51150"
                     /db_xref="SPTREMBL:Q01924"
                     /translation="MNNKMPLNKEAGFLAHTKRKRFAVTLVGVFFMLLASAGAIGFG
                     QVAYAADEKTVPHRVSQNPEPFWYGIDFYKGPYTRYHNLQLNLNGSKTYQAYCFNLKR
                     FEPKKEGSYFPNWKRMDGSSEETFVKYADNRKNESSRVIDVELEVELRNKLILRVLINGYP
                     NNGNIMEGLEPLNAILVTDNAVWYYSDNSSIFNTDNEFFTTEAKDLNIKPEQLSLMRV
                     ALKKLIDPKLSEESLKPVPSTFRLNIFESODKLYQNLLSAEFVPENPKPGETPEHGP
                     KTPELDGTPIPEGPQRPNESSLEPTLPPVMLDGQEVPEVPSESLEPALPLPLMPELDGQE
                     VPEVPSESLEPALPLMPELDGQEYPEKPSVDLPIEDPRYEFNNKDQSPLAGESCETE
                     YITEVYGVNGQNPVDIDKKLPNETGFSGNMVETEDTKEPGVLMGGQSESVEFTKDTQTG
                     MSGQTTPQVETEDTKEPGVLMGGQSESVEFTKDTQTGMSGQTTPQVETEDTKEPGVLM
                     GGQSESVEFTKDTQTGMSGQTTPQVETEDPKEPGVLMGGQSESVEFTKDTQTGMSGFS
                     ETVIIVEDTRPKLVFHFDNNEPKVEENREKPTKNITPILPATGDIENVLAFLGILILS
                     VLPIFSLLKKQTKQ"
BASE COUNT       790 a     383 c     477 g     613 t
ORIGIN
```

```
alignment_scores:
         Quality:    603.00                Length:      755
         Ratio:      1.519                 Gaps:         32
Percent Similarity:  52.583     Percent Identity:    27.417

alignment_block:
US-09-494-297-2 x SPSFB

Align seg 1/1   to: SPSFB     from: 1  to: 2263
```

```
  1  to: SPSFB   .
```

```
 20 SerLysAsnSerLysArgPheThrValThrLeuValGlyValPheLeuMe    36
    :::|||||||:||||||||||||||||||:||||||||||||||||:||
172 ACAAAAAGAAAAAAGGCGATTTGCTGTCACTTAGTGGGAGTCTTTTTTAT   221

 36 tIlePheAlaLeuValThrSerMetValGlyValAlaLysThrValPheGlyL  53
    |||:::||:|||:::|||||:::|||||||||:::|||:::||:|||:::
222 GCTTTTGCCATCCGCGGGTGCTATCCGTTTCGGTCAAGTAGCCCTATGCTG   271

 53 euValGluSerSerThrProAsnAlaIleAsnProAspSerSerSerGlu    69
    |||::|||:|::|||||||||||||||||||::|||||:::|
272 CCGATGAGAAGACTGTGCCCTCATAGAGTTAGT......CAAAATCCTAG   315

 70 TyrArgTrpTyrTyrGlyTyrValArgGlyHisProTyrTyrLy          86
    :::|||||||||||:::||||||||:::|||||:::||||:||:
316 TTTCCCTGGTATGGTTATGATTTCTAC...AAAGGA...CCTTACACAAG   359

 86 sGlnPheArgValAlaHisAspLeuArgValAsnLeuGluGlySerArgs    103
    ||:|||:::|:||||||:||||||::|||:||:::|||::|||:::||:
360 ATAT.........CATAATTTACAGCTAAATCTCAACGGAAGTAAA       397

103 erTyrArgGlnValTyrCysPheAsnLeuLysLysAlaPheProLeuGlySer  119
    |||:||||||::||:::||||:||||||||:::|||||||:||:::|||
398 CATATCAGGCCTATCGTCATTGTTTAATTTAAAGAGATTTGAACCTCAAGA   447

120 AspSerSerValLysTyrTrpTyrLysLysHisAspGlyIleSerThrLy    136
```

```
448 GGTAGTGTATTTTCCGGAATTGGTATAAGAGGTGGGATGGTAGTGAAGAGAC  497
    |||:::::||||||||||||||||||||:::||||||||||||||:::
136 sPheGluAspTyrAlaMetSerPro.............Argl           146
    ||||||                                  |||:
498 ATTTGTAAAATATGCTGATAACCCCAGAAAAGATAATGAAAGTTCCAGAG   547

146 leThrGlyAspGluLeuAsnGlnLysLeuArgAlaValMetTyrAsnGly    162
    :::||||||||||||||||||||||||:::||||||||||:::||||||
548 TTATAGAGATGTAGAATTAGAAAAAAATATATTATTAAGAGAGTGCTTAGGA  597

163 HisProGlnAsnAlaAsnGlyIleMetGluGlyLeuGluProLeuAsnAl    179
    ||||||:::|||||||:::|||||:::|||||||||||||||||:||||
598 TATCCGAATAACCGTAATGGAATCATGGAGGGTTTAGAACCATTAAATGC   647

179 aIleArgValThrGlnGluAlaValTrpTyrTyrSerAspAsnAlaProI    196
    :::|||||||:::||||||:::||||||||||||:::||||||:::|||
648 TATCTTAGTAACTCAAGAAGCTGTTTGGTACTATTCTGATAATTCATCTA   697

196 leSerAsnProAspGluSerPheLysArgGluSerGluSerAsnLeuVal    212
    |||:::||||||:::|||||||||:::|||||||:::|||||||:||||
698 TTTTTAATACTGATAACCTTCTTTACGACTGAAGCTAAAGATCTTAATATA  747

213 SerThrSerGlnLeuSerLeuMetArgGlnAlaLeuLeuLysGlnLeuIleAs  229
    :::||||||:::||||||||||||:::||||||:::|||||||:::|||
748 AAACCTGAGCAANTTATCATTAATGCGTGTTGCATTAAAGAAAATTGATAGA  797

229 pProAsnLeuAlaThrIleMetProLysGlnValProAspAspPheGlnLnI  246
    |||:::|||||||||::||||||:::|||||||:::||||||:::|||||
798 TCCTAAGTTATCTGAAGAATCATTGAAACCTGTGCCATCTACTTTTAGAC   847

246 euSerIlePheGluSerGluAspLysProLeuLysSerGlyLysTyr    262
    :::|||||||:::|||||||:::||||||:::||||||:::|||||||
848 TGAATATTTTTCGAATCTCAAGGAT..................AAGCTATAT   879

263 GlnAsnLeuLeuSerGlyLysGlyLeuValProThrLysProThrProGl    279
    |||||||||||:::|||||||:::||||||:::||||||:::||||||:::
880 CAAAATCTTTCTAAGTGCTAAGGTTTGTACCAGAAAATCCCCCTAAACCTGG  929

279 yAspProMetProProAsnGlnProGlnThrThrSerValLeuIleA    296
    |::|                                      :::|||
930 ACCCCTGAACATGGACCCAAAACTCCCGAGTTG.........            966

296 rAsnAspIleGlyAspValIleGluLeuSerAspGlyThrThrTyrThrLeuT  312
    :::                             |||||||  :::
967 ..................................GATGGCACTCCAATTCCCGAAGGCCCA  993

313 GlnLeuThrGlyAspAsnValAsnSerPheGlnAlaArgValPheSerSe    329
    :::::|                                       :::
994 CAACGTCCAAATGAGAGTTTAGAACCTACAGCGTTCCACCAGTG.......  1035

329 rAsnAspIleGlyArgIleGluLeuSerAspGlyThrThrTyrThrLeuT    346
    |||||||                                 |||||||
1036 ..........................................AIGCTAGACGGA...CAAGAAGTCC  1057

346 hrGluLeuAsnSerProAlaGlyThrSerIleAlaGluProIleThrPhe  362
    :::::|                                       |||:::
1058 CCGAAGTTCCAAGTAGAGAGCTTAGAACCTGCGCGTTCCCCCATTGATGCCA  1107

363 LysValGlyAlaGlyLysValGlyValTyrThrIle............    372
    |||||||:::||||||:::||||||:::
1108 GAGCTAGAATGGGCCAAGAAAGTCCCAGAAGCTTAGAACC           1157

373 ......IleAspGlyLysGlnIle...GluA    380
          :::||||||:::||||||   |||:
1158 TGGCGTTCCCCCATTGATGCCAGAGTTAGATGGTCAAGAAGTCCCCTGAAA  1207

380 snProAsnLysGlu...IleValGluProTyrSerValGluAlaTyr    394
    :::|||||:::      :::||||||:::||||||:::||||||
1208 AACCTAGTGTTGACTTACCTATTGAAGATCCTCGT............     1242

395 AsnAspPheGluGluPheSerValLeuThrThrGlnAsnThrTyrAlaLysPh  411
```

```
1242  .....................................................  1242

 411  eTyrTyrAlaLysAsnLysAsnGlySerGlnValValTyrCysPheA            428

1243  .....................................................  1243

 428  snAlaAspLeuLysSerProProAspSerGlnAspGlyGlyLysThrMet        444

1253  ACAATAAAGACCAATCACCTTCTAGCGGGTCAG                          1284
                 |||   ::|
 445  ThrProAspPheThrThrGlyGluValLysTyr...ThrHisIleAlaGl        460

1285  ...............TCTGGTGAGACGGAGTATATTACCGAAGTTTATGG         1319
                     ::||||||||:::||||||  |||:::::  ||
 460  yArgAspLeuPheLysTyrThrValLysProArgGlyAspThrAspAspT        477

1320  AAATCAA..........CAGAACCCTGTTGATATTGATAAA....             1350
      |||              ||||         |||
 477  hrPheLeuLysHisIleLysLysValIleGluLysGlyTyrArgGluLys        493

1351  ......................AAACTTCCGAAATGAAAACAGGTTTT......TCA  1377
                            ::|||                        |||
 494  GlyGlnAlaIleGluTyrSerGlyLeuThrGlnThrGlnLeuArgAlaAl        510

1378  GGAAATATGGTTGAGACACGAAGAATACGAAAGAGCCAGGAGTGTTGATGGG      1427
      |||
 510  aThrGlnLeuAlaIleTyrTyrPheThrAspSerAlaGluLeuAspLysA        527

1428  AGGTCAA.........................AGTGAGTCTGTTGAATTTACTAAAG  1459
      |||                            :::||||||||:::||||||
 527  spLysLeuLysAspTyrHisGlyPheGlyAspMetAsnAspSerThrLeu        543

1460  ACACTCAAACAGGCATGAGTGGT...............CAAACAACTCCT         1494
      |||
 544  AlaValAlaLysIleLeuValGlnTyrrAlaGlnAspSerAsnProPro.        559

1495  CAGGTTGAG................ACAGAAGACACGAAAAGAGCCAGG         1526
      |||                     :::
 560  ...........GlnLeuThrHisAspLeuAspPhePheIleProLeuAsnA       571

1527  AGTGCTGATGGGAGGTCAAAGTGAGTCTGTTGAATTACCAAAGATAACTC        1576
      |||
 571  snAsnLysTyrGlnSerLeurIleGlyThrGlnThrTrpHisProGluAspLeu    587

1577  AAACAGGCATGAGTGGTCAAACACGTTCTCAGGTTGAGACAGAAGATACG        1626
      :::
 588  ValAspIle...IleArgMetGluAspLysLysGlyGluValIleProValTh     603

1627  AAAGAGCCAGGAGTGCTGATGGGAGGCCAAAGTGAGTCTGTTGAATTTAC        1676
      |||
 603  rHisAsnLeuThrLeuArgLysThrValThrGlyLeuAlaGlyAspArgT        620

1677  T................AAAGACACTCAAACAGGCATGAGTGGTCAAACAA        1711
      |                ||||||::::||
 620  hrLysAspPheHisPheGluIleGluLeuLysAsnAsnLysGlnGluLeu        636

1712  CTCCT.............CAGGTTGAGACAGAAGACACGAAAGAGCCAGGA       1749
      :::
 637  LeuSerGlnThrHisValLysThrAspLysThrAsnLeuGluPhe...LysAs    652

1750  GTATTGATGGGAGGTCAAAGTGAA...............TCTGTTGAATTTACTAAAGA  1793
      ::
 652  pGlyLysValAlaThrIleAsnLeuLysHisGlyLeuSerLeuThrLeuGln      669

1794  CACTCAAACAGGCATGAGC...........GGTTTCAGTGAAACAGTGACCA       1834
      |||::::||
 669  lyLeuProGlyGlyTyrSerTyrLeuValLysGluThrAspSerGluGly        685

1835  TTGTTGAAGATACCCGTCCGAAGTTAGTTGTTCCAATTTGACAATAATGAG       1884
      :::
 686  TyrLysValLysValAsnSerGlnLeuValAlaAsnAlaThrValSerLy        702

1885  CCCAAAGTGGAAGAGAAGAATCGGGAAAAAGCCTACA                     1917
      ||| |||||:::
 702  sThrGlyIleThrSerAspGluThrLeuAlaPheGluAsnAsnLysGluP        719

1918  ..................................AAAAATATAACAC            1930
                                        |  |
 719  roValValPro...ThrGlyValValAspGlnLysSerIleLeuAsnGlyLeuTyrTyrLeuAla  734

1931  CTATCCTTCCTGCCAACAGGAGATATTGAGAAGTGTTTGGCCTTTCTTGGA       1980
      :::::||||:||  ||||||  :::::||||||:::::::||||::
 735  LeuIleValIleAla                                           739

1981  ATCCTTATTTTGTCA                                           1995
      :::::
```

seq_name: gb_ba2:GGU31115

seq_documentation_block:

FEATURES   Location/Qualifiers
     source   1..2066
                     /organism="Streptococcus sp. 'group G'"
                     /strain="1750"
                     /db_xref="taxon:1320"
     gene   252..1994
                     /gene="gfbA"
     CDS   252..1994
                     /gene="gfbA"
                     /note="GfbA"
                     /codon_start=1
                     /transl_table=11
                     /product="group G streptococcal fibronectin binding
                     protein"
                     /protein_id="AAB06623.1"
                     /db_xref="GI:950169"
                     /translation="MFLRGEKMNNKMFLNKEAGEVHHTKRRRRFAVTLVGVFFMLLACA
                     GAIGFGQVAYAAADEKTVPNFKSPDPDYPWYGYDAYTGAFLRYHDLNVNLEGSTPYQVY
                     CFNLVRQEFSKVNGFRPKFRFKKVDGDNAVFKQYAANPRVIDGDLERNILNVVINGYPS
                     DANGIMKGLDRYNAILVTQYAIWYSDNSKFIVDEQWKSELRDGEISESQVKLMRDAL
                     RKLISPDLEKTVKNQLPENYKLNLFISKDNSIQNLLSTEYVQDNLQKPEEPKEESPK
                     EENPKIPDLDGHEIPEVPQEPSDPVIPPVILDGEEVPEPSESLEPALPLMPELDGQ
                     EVPERPSIDLPIEVPRYEFNNKDQSPLAGESGETEYITEVYGNQONPVDIDKKLPNET
                     GFSGNMVETEDTKEPEVLMGGQSESVEFTKDTQIGMSGQTTPQVETEDTKEPGVLMGG
                     QSESVEFTKDTQTGMSGQTTPQVETEDTKEPGVLMGGQSESVGFTKDTQIGMSGFSET
                     ATVVEDTRPKLVFHFDNNEEPKVEENQEKPFKNITPILPAPGDIENVLAFLGILILSVL
                     PIFSLLKNQNNKY"
BASE COUNT   724 a   326 c   435 g   581 t
ORIGIN

alignment_scores:
     Quality: 523.00   Length: 728
     Ratio: 1.441   Gaps: 23
     Percent Similarity: 49.863   Percent Identity: 24.725

```
alignment_block:
US-09-494-297-2 x GGU31115

Align seg 1/1  to: GGU31115  from: 1  to: 2066

20   SerLysAsnSerLysArgPheThrValThrLeuValGlyValPheLeuMe    36
     ::|||:|||||::|||:||||||||||||||||||||||||||:|||:||
318  ACAAAAAGAGAAAAAAGGCGATTTCGTGTCACTTTAGTGGGAGTCTTTTTAT  367

36   tIlePheAlaLeuValThrSerMetValGlyAlaValLysThrValPheGlyL  53
     :||::|||::|||                        :||||
368  GCTTTTGCCATCCGCGGGGTGCTATCGGGTTTGGTCAAGTAGCCTATGCCTG  417

53   euValGluSerSerThrProAsnAlaIleAsnProAspSerSerSerGlu    69
     |||||||||:::||||||                                ::|
418  CCGATGAGAAGACTGTGCCGAAATTTAAAAGCCCAGAT.....CCAGAT    461

70   TyrArgTrpThrTyrGlyThrTyrGluSerTyrValArgGlyHisProTyrTyrLy  86
     |||||||||||||||:|||||||:|||||||||                 :||
462  TARCCCTGGTTATGGCTATGGCTATGCTGAGGCGTTTTAAGATAT..      509

86   sGlnPheArgValAlaHisAspLeuArgValAsnLeuGluGlySerArgS    103
                                      .CATGATTTAAAATGTTA    510

103  ertyrArgGlnValThrTyrCysPheAsnLeuLysLysAlaPheProLeuGlySer  119
     ||||||||||                       :||||||||
544  CTTACCAAGTTTATTGCTTTAATTTAGTAAGACAGGAACCGAGTAAGGTT    593

120  AspSerSerValLysLysTrpTyrLysLysHisAspGlyIleSerThrGly    136
     |||||:::                                :||||||||||
594  AATGGTTTCAGAAAATTCCGGTTCAAAAAAGTTGATGGTGATAATGCTGT    643

136  sPheGluAspTyrAlaMetSerProArgIleThrGlyAspGluLeuAsn    153
     |||:|||                        :||||||:||
644  GTTTAAACAATATGCAGCTAATCCTAGAGTAATTGATGGAGAATTTAGAAC  693

153  lnLysLeuArgAlaValMetTyrAsnGlyHisProGlnAsnAlaAsnGly    169
     ||:::||||||||                   ::||||:||
694  GTAACATTTTAAAATGTAATTTACAATGATATCCAAAGTTGCAAATGGG    743

170  IleMetGluGlyLeuGluProLeuAsnAlaIleAlaArgValThrGlnVal    186
     ||||||||||||                     :|||||||
744  ATTATGGAAGGATTGGACCGGTACAATCCTATTTTAGTRAACTCAAACTGC  793

186  aValTrpThrTyrTyrSerAspAsnAlaProIleSerAsnProAspGluSerP    203
     :|||||||||||||||||||||||             :||
794  GATTTGGTATTATTCTGATAACTCAAAGTTATA..GTGGATGAACAGT    840

203  heLysArgGlySerGluSerAsnLeuValSerThrSerGlnLeuSerLeu    219
     ::|||                        |||||||||||::||||
841  GGAAGTCTGAATTAAAGAGATGGAGAAAATTAGCCAAAGTCAAGTAAGTTA  890

220  MetArgGlnAlaLeuLysGlnLeuIleLeuAspProAsnLeuAlaThrLysMe    236
     |||||||||||||||                     |||||||||
891  ATGCGTCAGGCTCTTAGAAAATTAATCACCAGATTTAGAAAAAACAGT    940

236  tProLysGlnValProAspAspPheGlnLeuSerIlePheGluSerGluA    253
     :|||||||||||||                     |||::|
941  CAAAAAACAGTTGCCTGAGAATTATAAGTAAATCTTTTTATTCAAAAG    990

253  spLysGlyAspLysLysTyrAsnLysGlyTyrGlnAsnLeuLeuSerGlyGly  269
     :|||                         ||||||
991  AT.                        AACTCGATTCAAAATTTATTGAGTACTGAG  1022

270  LeuValProThrLysProProGlyProProGlyProMetProProAspS    286
     :||:                        |||||||
1023 TATGTACAAGATAACCTCCAAAAAACCAGGTGAAGAACCT.    1061

286  nGlnProGlnThrThrSerValLeuLeuIleArgLysTyrAlaIleGlyAspT    303
     ..                       :||||
1061 .................    1061
```

```
303  yrSerLysLeuLeuGluGlyAlaThrLeuGlnLeuThrGlyAspAsnVal    319
                                      .................    1061

320  AsnSerPheGlnAlaArgValPheSerSerAsnAspIleGlyGluArgIl    336
     ||| :||::::|||                      :|||
1062 .......AAAGAGGACGAGTCCAAAAAGAAGAGAACCCCAAAAT    1096

336  eGluLeuSerAspGlyThrTyrThrLeuTrpThrGluLeuAsnSerProAlaG    353
     ::|||                        |||||||||||:
1097 TCCAGAATTTAGATGGG...CATGAAAATCCCAGAAGTTCACACAGAAC  1143

353  lyTyrSerIleAlaGluProIleThrPheLysValGluAlaGlyLysVal    369
     :|||                        ||||||||||
1144 CAGACCCAGTTATCCCACCGGT.   1166

370  TyrThrThrIleLeuAspGlyLysGlnIleGluAsnProAsnLysGluLeuVa    386
     |||||:||||||||||||||||:||:|||                 :||
1167 .......ATTTTGGATGGTGAAGAAGTCCCAGAAGTTCCAAGCAGAGCTT  1210

386  lGluProTyr.                       SerValGluAlaTyrAsnAspP    397
     |:|||                         ::|
1211 AGAACCTGCGGCTTCCTCCCTCCATTAATGCCACAGAGTTAGATGGTCA     1260

397  heGluGlnPheSerValLeuThrThrGlnAsnThrValAlaLysPheTyrTyr    413
     ||||:|||                         :||
1261 CTGAAAAACCTAGTTAGTTGACTTACCTATTGACAGGTTCCTCGT.   1301

414  AlaLysAsnLysAsnGlySerSerGlnValValTyrCysPheAsnAlaAs    430
     |||||||                         ||||||||||
1302 AGAACCTGCGGCTTCCTCCCTCCATTAATGCC.     TATGAGTTTAACAATAA  1318

430  pLeuLysSerProProAspSerGlyGlyGlyLysThrMetThrProAro    447
     :|||                         ||||||||||
1319 AGACCAAATCGCCTCTAGCGGGTGAG.    1343

447  spPheThrThrGlyValValLysTyr.     ThrHisIleAlaGlyArgAsp    462
     :|||||||                          |||
1344 .......TCTGGTGACAGAAGAGATACCAAAGAGTGTTGATGGGAGGCCAA  1385

463  LeuPheLysThrTyrThrValLysSerProArgAspThrProAspThrPhe    479
     ::|||                         |||
1386 .................    CAGAACCCTGTTGATATTGATAAA.   1409

479  uLysHisIleLeuLysValIleGluLysGlyTyrArgGluLysGlyGln    496
     |||                         ::|
1410 .......AAACTTCCGAATGAAAACAGGTTTT.       TCAGGAAATA  1443

496  laIleGluTyrSerGlyLeuThrGluThrGlnLeuGluArgAlaAlaThrGln    512
     ::|||||                          :||
1444 TGGTTGACAGAAGAGATACCAAAGAGTGTTGATGGGAGGCCAA    1493

513  LeuAlaIleGluTyrTyrPheThrAspSerAlaGluLeuAspLysLysLysLe    529
     :::||||:|||                      ||||
1494 .......AGTGAGTCTGTTGAATTTACGAAAGACACTCA     1525

529  uLysAspTyrHisGlyPheGlyAspMetAsnAspSerThrLeuAlaValA    546
     |||                        ::|
1526 AACAGGCATGAGTGGT.     1541

546  laLysIleLeuValGluTyrAlaGlnAspSerAsnProProGlnLeuThr    562
     :::|||                        |||
1542 CAAACAACTCCTCACGGTTGAG     1562

563  AspLeuAspPhePheIleProAsnAsnAsnLysTyrGlnSerLeuIleGl    579
     ||||::|                         :||
1563 ACAGAAGAGATACGAAAAGAGCCAGGA.    1586

579  yThrGlnTrpHisProGluAspLeuValAspIleIleArgMetGluAspL    596
     |||                        :||
1587 .................    GTGTTGATGGGAGGTC     1602

596  ysLysGluValIleProValThrHisAsnLeuThrLeuArgLysThrVal    612
```

```
1603 AAAGTGAGTGTCTGTTGAATTACG..............AAAGATACTCAA 1637
      :::::|||                                  |||
 613 ThrGlyLeuAlaGlyAspArgThrLysAspPheHisPheGlnIleGlu 629
     ||:||:|:::|||::||:||
1638 ACAGGCAATGAGTGGTCAAACACTCCT...........CAGGTTGACAG 1675
      |||||:::|||::|:||
 629 GlyAsnGluSerGlnThrValLysThrAspLysT 646
     :::|||:||||:|||||::||
1676 AGAAGATACGAAAGAGCCCAGGAGAGTGTTGATGGGAGGTCAAAGTGAGTCTG 1725
      :::||:||||||::||::|
 646 HisAsnLeuGluPheLysAspGlyLysAlaThrIleAsnLeuLysHisGly 662
     ::|||:||:::||
1726 TTGGATTTACG..AAAGACACTCAAACAGGCATGAGT..GGT 1763
      |||:||                    ||:||            |||
 663 GluSerLeuThrLeuGlnGlyLeuProGlyLeuGlyTyrSerTyrLeuValLy 679
     ||:|||   |||                          ||:|||
1764 TTCAGTGAAACAGCGACTGTGTTGTGAAGATACGCGTCCGAAGTTAGTGTT 1813
      |||   |||                                 ||:|||
 679 sGluThrAspSerGluGlyTyrLysValLysValAlaSerGlnGluValA 696
     :::||||||                          :::||
1814 CCATTTTGACAATAATGAGCCCAAAGTGGAAGAGAATCAGGAAAAAACCTA 1863
      :::||||||                          :::||
 696 laAsnAlaThrValSerLysThrGlyIleThrSerAspGluThrIleuAla 712
     :::|||:::||||||:::|||
1864 CA............................................. 1865
 713 PheGluAsnAsnLysGluProValValPro...ThrGlyValAspGlyIle 728
     ||:|||   |||                             :||
1866 ......AAAAATAACACCTARCCTTCCTGCAACAGAGAATATTGAGAA 1909
      ::||||   |||                               :||
 728 sIleAsnGlyTyrTyrLeuAlaLeuTleValIleeAla 739
     :::||||:::|||||:::||||:::||||:
1910 TGTTTGGCCTTTCTTGGAATCCTTATTTGTCA 1943
```

```
seq_name: gb_bal:AF009914

seq_documentation_block:
LOCUS       AF009914           321 bp    DNA            BCT       13-FEB-1998
DEFINITION  Streptococcus pyogenes fibronectin binding protein F gene, partial
            cds.
ACCESSION   AF009914
VERSION     AF009914.1  GI:2267181
KEYWORDS    .
SOURCE      Streptococcus pyogenes.
  ORGANISM  Streptococcus pyogenes
            Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
            Streptococcus.
REFERENCE   1  (bases 1 to 321)
  AUTHORS   Katerov,V., Andreev,A., Schalen,C. and Totolian,A.A.
  TITLE     Protein F, a fibronectin-binding protein of Streptococcus pyogenes,
            also binds human fibrinogen: isolation of the protein and mapping
            of the binding region
  JOURNAL   Microbiology 144 (Pt 1), 119-126 (1998)
  MEDLINE   98129085
REFERENCE   2  (bases 1 to 321)
  AUTHORS   Katerov,V.E., Andreev,A.S., Schalen,C. and Totolian,A.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-1997) Medical Microbiology, Lund University,
            Solvegatan 23, Lund 22362, Sweden
FEATURES             Location/Qualifiers
     source          1..321
                     /organism="Streptococcus pyogenes"
                     /strain="M4 strain Umea"
                     /db_xref="taxon:1314"
     CDS             <1..>321
                     /codon_start=2
                     /transl_table=11
                     /product="fibronectin binding protein F"
                     /protein_id="AAC38161.1"
                     /db_xref="GI:2267182"
                     /translation="FCQVAYSADEKTVPNFKSPDDYPWYGYYAYAGKGYPGYDISKYY
                     HDLRVNLNGSQVYQVCFNIQKIFPYNVKSVTQKWFKKAEGNSDTFGPYAMNRVQGE"
```

```
BASE COUNT       107 a       48 c       68 g       98 t
ORIGIN
```

```
alignment_scores:
              Quality:  272.50             Length:     108
                Ratio:    3.633            Gaps:         2
   Percent Similarity:   69.444   Percent Identity:   49.074
```

```
alignment_block:
US-09-494-297-2 x AF009914

Align seq 1/1  to: AF009914   from: 1  to: 321
```

```
  1 PheGlyLeuVal.................GluSerSerThrProAsnAlaIl 62
    ||||||:::                      |||||:::::|||||:|
  2 TTTGGTCAAGTAGCCTATTCTGCCGGATGAGAAAGACGTGCCGAATTTAA 51
```

```
 51 sLysAlaPheProLeuGlySerAspSerSerValLysLysTrpPheLysL 129
    :::|||:|||:::||                      ||:||||:::|
 52 AAGCCCAGAT......CCAGATTATCCCTGGTATGGTTATTATGGCTATAG 95
```

```
 62 eAsnProAspSerSerGluTyrArgThrProTyrGlyTyrGlnSerTyrrV 79
    :::|||:|||                        ||||||:::|||
  2 AAGCCCAGAT......CCAGATTATCCCTGGTATGGTTATTATGGCTATAG 95
```

```
 79 aIleArgGlyHisProTyrTyrLysGlnPheArgValAlaHisAspLeuArg 95
    :::||:||||:::||                      ||:|||
 96 GTAAGGGGTATCCCTGGATATGATATCAGTAAGTATTATCATGATTTAAGG 145
```

```
 96 ValAsnLeuGluGlySerArgSerTyrGlnValTyrCysPheAsnLeuTyr 112
    |||:::||||:||                        |||:::|||
146 GTAAATCTAAACGGAAGTCAGGTCTACCAAGTTTATTGCTTTAATATTCA 195
```

```
112 sLysAlaPheProLeuGlySerAspSerSerValLysLysTrpPheLysL 129
    ||:||||:::||                        |||:::||||:
146 GTAAATCTAAACGGAAGTCAGGTCTACCAAGTTTATTGCTTTAATATTCA 195
```

```
129 ysHisAspSerGlyIleSerThrLysPheGluAspTyrAlaMetSerProArg 145
    ||:::|||                        |||:::||||:::|
246 AAGCCGAGGGGAATTCTGATACTTTTGGACCTTATGCAAATGAATCCAAGA 295
```

```
146 IleThrGlyAspGluGluAsnGln 153
    ||:::|||
296 GTCCAAGGAGAGAAGAACTTAGTCAA 319
```

```
seq_name: gb_bal:AF009913

seq_documentation_block:
LOCUS       AF009913           313 bp    DNA            BCT       13-FEB-1998
DEFINITION  Streptococcus pyogenes fibronectin binding protein F gene, partial
            cds.
ACCESSION   AF009913
VERSION     AF009913.1  GI:2267179
KEYWORDS    .
SOURCE      Streptococcus pyogenes.
  ORGANISM  Streptococcus pyogenes
            Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
            Streptococcus.
REFERENCE   1  (bases 1 to 313)
  AUTHORS   Katerov,V., Andreev,A., Schalen,C. and Totolian,A.A.
  TITLE     Protein F, a fibronectin-binding protein of Streptococcus pyogenes,
            also binds human fibrinogen: isolation of the protein and mapping
            of the binding region
  JOURNAL   Microbiology 144 (Pt 1), 119-126 (1998)
  MEDLINE   98129085
REFERENCE   2  (bases 1 to 313)
  AUTHORS   Katerov,V.E., Andreev,A.S., Schalen,C. and Totolian,A.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-1997) Medical Microbiology, Lund University,
            Solvegatan 23, Lund 22362, Sweden
FEATURES             Location/Qualifiers
     source          1..313
                     /organism="Streptococcus pyogenes"
                     /strain="M4 strain 281C"
                     /db_xref="taxon:1314"
```

Left column

```
CDS             <1..>313
                /codon_start=2
                /transl_table=11
                /product="fibronectin binding protein F"
                /protein_id="AAC38160.1"
                /db_xref="GI:2267180"
                /translation="FGQVAYSADEKIVPNFKSPDDYPWGYDAYGKGYPGYDISKYY
                HDLRVNLNGSQVYQVYCFNIQKIFPYNVKSVTQKWFKKAEGNSDTFGLHAMNPRVGGE
                EL"

BASE COUNT       102 a     47 c     68 g     96 t
ORIGIN

alignment_scores:
   Quality:   262.50          Length:    106
   Ratio:     3.547           Gaps:      2
   Percent Similarity: 69.811      Percent Identity: 48.113

alignment_block:
US-09-494-297-2 x AF009913

Align seg 1/1  to: AF009913   from: 1  to: 313  .

51 PheGlyLeuVal..............GluSerSerThrProAsnAlaIl 62
   ||||||  |||              ||  ||::||:|||||||||||
2 TTTGGTCAAGTAGCCTAATTCTCGGATGTGAGAAGACTGTGCCGAATTTTAA 51

62 eAsnProAspSerSerGluTyrArgtrpTyrGlyTyrGlyLuSerTyrV 79
   ::||:|||              ||||||||||||||||||::||||||||||
52 AAGCCCAGAT.......CCAGATTATCCCTGGTATGGTTATGGTTATG 95

79 aArgGlyHisProTyrTyrLysGlnPheArgValAlaHisAspLeuArg 95
   :||||||||||||||||||||| ||||||||||||||||||:|||||||
96 GTAAGGGGTATCCTGGATATGATAAAGTTAGGTATATCATGATTTAAGG 145

96 ValAsnLeuGluGlySerArgSerTyrGlnValTyrCysPheAsnLeuY 112
   ||||||||||||||||||||||::||||||||||||||||||||||||
146 GTAAATCTAAACGGAAGTCAGGTCTACCAAGTTTATTGCTTTAATATTCA 195

112 sLysAlaPheProLeuGlySerAspSerSerValLysLysTrpTyrLysL 129
   ::| ||  ||:||||||||:||||||||||||:::||||||||||||||
196 AAAAAATTTTCCGTACAATGTGAAAAGCGTTACTCAAAAAATGGTTTAAA 245

129 ysHisAspGlyIleSerThrLysPheGluAspTyrArgAlaMetSerProArg 145
   :||||||||||||||::|||||              ::||||||||||
246 AAGCCGAGGGGAATTCGATACTTTGGACTTCATGCAAATGAATCCAAGA 295

146 IleThrGlyAspGluLeu 151
   ::|||||||||||||||
296 GTCCAAGGAGGAGAACTT 313
```

seq_documentation_block:

seq_name: gb_bal:AF009910

```
LOCUS       AF009910      326 bp    DNA        BCT     13-FEB-1998
DEFINITION  Streptococcus pyogenes fibronectin binding protein F gene, partial
            cds.
ACCESSION   AF009910
VERSION     AF009910.1  GI:2267173
KEYWORDS    .
SOURCE      Streptococcus pyogenes.
  ORGANISM  Streptococcus pyogenes
            Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
            Streptococcus.
REFERENCE   1  (bases 1 to 326)
  AUTHORS   Katerov,V., Andreev,A., Schalen,C. and Totolian,A.A.
  TITLE     Protein F, a fibronectin-binding protein of Streptococcus pyogenes,
            also binds human fibrinogen: isolation of the protein and mapping
            of the binding region
  JOURNAL   Microbiology 144 (Pt 1), 119-126 (1998)
  MEDLINE   98129085
REFERENCE   2  (bases 1 to 326)
  AUTHORS   Katerov,V.E., Andreev,A.S., Schalen,C. and Totolian,A.A.
```

Right column

```
TITLE       Direct Submission
JOURNAL     Submitted (23-JUN-1997) Medical Microbiology, Lund University,
            Solvegatan 23, Lund 22362, Sweden
FEATURES             Location/Qualifiers
     source          1..326
                     /organism="Streptococcus pyogenes"
                     /strain="M4 strain M4/118"
                     /db_xref="taxon:1314"
     CDS             <1..>326
                     /codon_start=2
                     /transl_table=11
                     /product="fibronectin binding protein F"
                     /protein_id="AAC38157.1"
                     /db_xref="GI:2267174"
                     /translation="FGQVAYAADEKTVPNRINSNPEPPWYGYDSYSGGRLLRYHNLKVN
                     LNGSKEYQAYCFNLKRFEPKKEESSSPNWYKKLDGSTETFKKYAENPRFSGEELRHI
                     LKVLYM"

BASE COUNT       118 a     41 c     75 g     92 t
ORIGIN

alignment_scores:
   Quality:   241.50          Length:    115
   Ratio:     3.136           Gaps:      3
   Percent Similarity: 66.957      Percent Identity: 44.348

alignment_block:
US-09-494-297-2 x AF009910

Align seg 1/1  to: AF009910   from: 1  to: 326  .

51 PheGlyLeuVal..............GluSerSerThrProAsnAlaIl 62
   ||||||  |||              ||::||::||||||||||:||||
2 TTTGGTCAAGTAGCCTATTCCAGCGGATGAGAAGACTGTGCCGAATTTTAA 51

62 eAsnProAspSerSerGluTyrArgTrpTyrGlyTyrGlyLuSerTyrV 79
   |||::|||              |||:||||||||||||||||||||||
52 TAAT.......TCAAATCCTGAGTTTCCTTGGTATGGTTATGGTTATA 95

79 aArgGlyHisProTyrTyrLysGlnPheArgValAlaHisAspLeuArg 95
   :||||||||||||||||||||||||              ::||||||
96 GCGGAAGATTGTTAAGATAT..............CACAATCTGAAA 127

96 ValAsnLeuGluGlySerArgSerTyrGlnValTyrCysPheAsnLeuY 112
   ||||||||||::||||||||||||::|||||||||||||||||||||
128 GTAAATCTAAATGGAAGTAAGGAGTATATCAAGCGTTATTGTTTTAACCTGAA 177

112 sLysAlaPheProLeuGlySerAspSerSerValLysLysTrpTyrLysL 129
   :| ||  ||:||||||||||||||||||||||||||||||||||||||
178 AAGGTTTGAGCCAAAAAAAGAAGAAAGTTAGCACAGTCCTAATTGGTATAAA 227

129 ysHisAspGlyIleSerThrLysPheGluAspTyrArgAlaMetSerProArg 145
   || ||||||||||||||::|||||||::||||||||:|||||||||||
228 AATTAGATGGGAGTACAGAAAACTTTTAAAAAAATACGCTGAAAAACCCTAGG 277

146 IleThrGlyAspGluLeuAsnGlyLysLeuArgAlaValMetTyr 160
   ::||||||||||||::||||::|||||||||||||||||||||
278 TTTAGTGGAGAGAATTAAGACGACATATATTGAAAGGTTCTCTAC 322
```

seq_documentation_block:

seq_name: gb_bal:AF009920

```
LOCUS       AF009920      348 bp    DNA        BCT     13-FEB-1998
DEFINITION  Streptococcus pyogenes fibronectin binding protein F gene, partial
            cds.
ACCESSION   AF009920
VERSION     AF009920.1  GI:2267193
KEYWORDS    .
SOURCE      Streptococcus pyogenes.
  ORGANISM  Streptococcus pyogenes
            Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
            Streptococcus.
REFERENCE   1  (bases 1 to 348)
```

(Left column)

```
AUTHORS    Katerov,V., Andreev,A., Schalen,C. and Totolian,A.A.
TITLE      Protein F, a fibronectin-binding protein of Streptococcus pyogenes,
           also binds human fibrinogen: isolation of the protein and mapping
           of the binding region
JOURNAL    Microbiology 144 (Pt 1), 119-126 (1998)
MEDLINE    98129085
REFERENCE  2  (bases 1 to 348)
  AUTHORS  Katerov,V.E., Andreev,A.S., Schalen,C. and Totolian,A.A.
  TITLE    Direct Submission
  JOURNAL  Submitted (23-JUN-1997) Medical Microbiology, Lund University,
           Solvegatan 23, Lund 22362, Sweden
FEATURES             Location/Qualifiers
     source          1..348
                     /organism="Streptococcus pyogenes"
                     /strain="M18 strain SS-36"
                     /db_xref="taxon:1314"
     CDS             <1..>348
                     /codon_start=2
                     /transl_table=11
                     /product="fibronectin binding protein F"
                     /protein_id="AAC38167.1"
                     /db_xref="GI:2267194"
                     /translation="FGQVAYAAADEKTVPNRINSNPEFPWYGYDAYTGAFLRYHDLNVN
                     LEGSTPYQVYCFNLVRQEPESKVNGFRKNWFKKVDGDNAVERKKYAANPRVIDRRFRTNI
                     LNVINGYSRVQMG"
BASE COUNT    115 a   45 c   78 g   110 t
ORIGIN

alignment_block:
US-09-494-297-2 x AF009920

Percent Similarity: 63.559     Percent Identity: 40.678
alignment_scores:
     Quality:  222.50          Length:  118
     Ratio:    2.967           Gaps:      3

Align seg 1/1   to: AF009920   from: 1  to: 348

  51 PheGlyLeuVal..............GluSerSerThrProAsnAlaIl  62
     ||||||    |||
   2 TTTGGTCAAGTAGCCTATGCTGCGGATGAGAAGACTGTGCCGAATTTTAA  51

  62 eAsnProAspSerSerGluTyrArgTrpTyrGlyTyrGluSerTyrV  79
     ::||||||||                |||||||||||||:::||||:
  52 AAGCCCAGAT.......CCAGATATATCCCTGGTACGGCTATGAGCTTATA  95

  79 aIArgGlyHisProTyrTyrLysGlnPheArgValAlaHisAspLeuArg  95
     ::   |||||
  96 CTGGAGCGTTTTTAAGATAT..............CATGATTAAAT  127

  96 ValAsnLeuGluGlySerArgSerTyrGlnValTyrCysPheAsnLeuV  112
     |||||||||||||||    ||||    ||||||||||||||
 128 GTTAATTTGGAAGGTAGCACGCCTACCAAGTTTATTGCTTTAATTTAGT  177

 112 sLysAlaPheProLeuGlySerAspSerSerValLysLysTrpTyrLysL  129
     |||                  :||||||||::||:::||||:
 178 AAGACAGGAACCGAGTAAGGTTTAATGGTTTAGAAAAAATTGGTTCAAA  227

 129 ysHisAspProGlyIleSerThrLysPheGluAspTyrAlaMetSerProArg  145
     ||                 ||||||   :::||||||
 228 AAGTTGAVGGTGATAAAGCTGTGTTTAAAAAAATATGCAGCTAATCCTAGA  277

 146 IleThrHisGlyAspGluLeuAsnGlnLysLeuArgAlaValMetTyrAsnGl  162
     :::||||||                  :::||||
 278 GTAATTGATCGGAGATTTAGAACGAACATTTTAAAATGTAATTACACACGG  327

 162 yHis  163
     |.::
 328 ATAT  331

seq_name: gb_ba1:AF009917
```

(Right column)

```
seq_documentation_block:
LOCUS       AF009917      278 bp    DNA             BCT       13-FEB-1998
DEFINITION  Streptococcus pyogenes fibronectin binding protein F gene, partial
            cds.
ACCESSION   AF009917
VERSION     AF009917.1  GI:2267187
KEYWORDS    .
SOURCE      Streptococcus pyogenes.
  ORGANISM  Streptococcus pyogenes
            Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
            Streptococcus.
REFERENCE   1  (bases 1 to 278)
  AUTHORS   Katerov,V., Andreev,A., Schalen,C. and Totolian,A.A.
  TITLE     Protein F, a fibronectin-binding protein of Streptococcus pyogenes,
            also binds human fibrinogen: isolation of the protein and mapping
            of the binding region
  JOURNAL   Microbiology 144 (Pt 1), 119-126 (1998)
  MEDLINE   98129085
REFERENCE   2  (bases 1 to 278)
  AUTHORS   Katerov,V.E., Andreev,A.S., Schalen,C. and Totolian,A.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-1997) Medical Microbiology, Lund University,
            Solvegatan 23, Lund 22362, Sweden
FEATURES             Location/Qualifiers
     source          1..278
                     /organism="Streptococcus pyogenes"
                     /strain="M15 strain 58/50"
                     /db_xref="taxon:1314"
     CDS             <1..>278
                     /codon_start=2
                     /transl_table=11
                     /product="fibronectin binding protein F"
                     /protein_id="AAC38164.1"
                     /db_xref="GI:2267188"
                     /translation="FGQVAYAAADEKTVPNRINSNPEFPWYGYDAYKGFEARYHDLKVN
                     LKGSKEYQVYCFNLKRSFPRRTHSITNNFYKKIVGSGSVFKSYAENPR"
BASE COUNT     89 a   38 c   59 g   92 t
ORIGIN

alignment_block:
US-09-494-297-2 x AF009917

Percent Similarity: 69.000     Percent Identity: 45.000
alignment_scores:
     Quality:  211.50          Length:  100
     Ratio:    3.065           Gaps:      3

Align seg 1/1   to: AF009917   from: 1  to: 278

  51 PheGlyLeuVal..............GluSerSerThrProAsnAlaIl  62
     ||||||    |||
   2 TTTGGTCAAGTAGCCTATGCTGCGGATGAGAAGACTGTGCCGAATCGTAT  51

  62 eAsnProAspSerSerGluTyrArgTrpTyrGlyTyrGluSerTyrV  79
     :||||||                  ||||||||||||:||:
  52 TAAT......TCAAATCCTGAATTCCTTGGTATGGTTATAGTGCTTATA  95

  79 aIArgGlyHisProTyrTyrLysGlnPheArgValAlaHisAspLeuArg  95
     ::|   |||
  96 AAGGATTTGAGGCAAGGTAT..............CATGATCTTAAA  127

  96 ValAsnLeuGluGlySerArgSerTyrGlnValTyrCysPheAsnLeuV  112
     |||||||||:::|||||||||||||||||||
 128 GTCAATTTAAAGGGAGTAAGGAATATATCAAGTTTATTGTTTTAAATTTAA  177

 112 sLysAlaPheProLeuGlySerAspSerSerValLysLysTrpTyrLysL  129
     |||                  |||:::||||||||||||:
 178 AAGAAGCTTCCTCGCCCGCACTCATAGTATTACAAATAATTTTACAGA  227

 129 ysHisAspGlyIleSerThrLysPheGluAspTyrAlaMetSerProArg  145
     ||                 :::|||||||||||:||||
 228 AAATTGTTGGGAGTGGATCCAGTGTTCAAAATCTTATGCAGAGAATCCTAGG  277
```

Left column

```
seq_name: gb_bal:AF009911
LOCUS       AF009911     310 bp    DNA             BCT       13-FEB-1998
DEFINITION  Streptococcus pyogenes fibronectin binding protein F gene, partial
            cds.
ACCESSION   AF009911
VERSION     AF009911.1  GI:2267175
KEYWORDS    .
SOURCE      Streptococcus pyogenes.
  ORGANISM  Streptococcus pyogenes
            Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
            Streptococcus.
REFERENCE   1  (bases 1 to 310)
  AUTHORS   Katerov,V., Andreev,A., Schalen,C. and Totolian,A.A.
  TITLE     Protein F, a fibronectin-binding protein of Streptococcus pyogenes,
            also binds human fibrinogen: Isolation of the protein and mapping
            of the binding region
  JOURNAL   Microbiology 144 (Pt 1), 119-126 (1998)
  MEDLINE   98129085
REFERENCE   2  (bases 1 to 310)
  AUTHORS   Katerov,V.E., Andreev,A.S., Schalen,C. and Totolian,A.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-1997) Medical Microbiology, Lund University,
            Solvegatan 23, Lund 22362, Sweden
FEATURES             Location/Qualifiers
     source          1..310
                     /organism="Streptococcus pyogenes"
                     /strain="M12 strain 100085"
                     /db_xref="taxon:1314"
     CDS             <1..>310
                     /codon_start=2
                     /transl_table=11
                     /product="fibronectin binding protein F"
                     /protein_id="AAC38158.1"
                     /db_xref="GI:2267176"
                     /translation="FGQVAYSADEKTVPNFKSPDDYPWYGYNSYRGIFARYHNLKVN
                     LKGSKEYQAYCFNLTKYFPRPTYSTNNFYKKIDGSGSAFKSYTANPRVLDENLDKLE
                     K"
BASE COUNT      112 a     47 c     58 g     93 t
ORIGIN
```

alignment_block:
US-09-494-297-2 x AF009911

alignment_scores:
```
      Quality:   211.50                    Length:    108
      Ratio:       2.820                    Gaps:        4
Percent Similarity:  69.444     Percent Identity:  42.593
```

Align seg 1/1 to: AF009911 from: 1 to: 310

```
  51 PheGlyLeuVal................GluSerSerThrProAsnAlaIle 62
     |||||| ||                   |||:|::||||||||||:||||||
   2 TTTGGTCAAGTAGCCTATTCTGCGGATGAGAGAGACTGTGCCGAATTTAA 51

  62 eAsnProAspSerSerSerGluTyrArgTrpTyrArgGlyTyrGluSerTyrV 79
     ::|||||||| |||:|||||||||||||||||||||||||||:||||||
  52 AAGCCCAGAT......CCAGATTATCCCTGGTATGGTTATAATTCGTAT. 94

  79 alArgGlyHisProTyrTyrLysGlnPheArgValAlaHisAspLeuArg 95
     |||                 |||||||||||||||||||||||||||||
  95 ..AGAGGA..........ATATTGCAAGAATATCACAATTTAAA 127

  96 ValAsnLeuGlyLysSerArgSerTyrGlnLnvalTyrCysPheAsnLeuLy 112
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 128 GTAAATCTAAAAGGAAGTAGAGTATCAAGGCGTATTGTTTTAACCTAAC 177

 112 sLysAlaPheProLeuGlySerAspSerSerValLysLysTrpTyrLysL 129
     |||                 ||||:|||:||::::::::||||||||||
 178 AAAATACTTCCTCCGCCCACTTATAGTACTACAAATAATTTTACAGA 227
```

Right column

```
seq_name: gb_bal:AF009918
LOCUS       AF009918     309 bp    DNA             BCT       13-FEB-1998
DEFINITION  Streptococcus pyogenes fibronectin binding protein F gene, partial
            cds.
ACCESSION   AF009918
VERSION     AF009918.1  GI:2267189
KEYWORDS    .
SOURCE      Streptococcus pyogenes.
  ORGANISM  Streptococcus pyogenes
            Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
            Streptococcus.
REFERENCE   1  (bases 1 to 309)
  AUTHORS   Katerov,V., Andreev,A., Schalen,C. and Totolian,A.A.
  TITLE     Protein F, a fibronectin-binding protein of Streptococcus pyogenes,
            also binds human fibrinogen: Isolation of the protein and mapping
            of the binding region
  JOURNAL   Microbiology 144 (Pt 1), 119-126 (1998)
  MEDLINE   98129085
REFERENCE   2  (bases 1 to 309)
  AUTHORS   Katerov,V.E., Andreev,A.S., Schalen,C. and Totolian,A.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUN-1997) Medical Microbiology, Lund University,
            Solvegatan 23, Lund 22362, Sweden
FEATURES             Location/Qualifiers
     source          1..309
                     /organism="Streptococcus pyogenes"
                     /strain="M18 strain M18/104"
                     /db_xref="taxon:1314"
     CDS             <1..>309
                     /codon_start=2
                     /transl_table=11
                     /product="fibronectin binding protein F"
                     /protein_id="AAC38165.1"
                     /db_xref="GI:2267190"
                     /translation="FGQVAYAADEKTVPNYSSPNPEFPWYGYDAYRGSFLRYHDLNVN
                     LEGSTSQVYCFNLVRQEPSKVNGLRKNWFFKKVDGDNAVFKKYAANPRVIDGDLERNI
                     "
BASE COUNT       98 a     38 c     69 g    104 t
ORIGIN
```

alignment_block:
US-09-494-297-2 x AF009918

alignment_scores:
```
      Quality:   210.50                    Length:    110
      Ratio:       2.924                    Gaps:        4
Percent Similarity:  65.455     Percent Identity:  42.727
```

Align seg 1/1 to: AF009918 from: 1 to: 309

```
 129 ysHisAspGlyIleSerThrLysPheGluAspTyrAlaMetSerProArg 145
     ||||||                 :::::|||::||||||||||:||||||
 228 AAATTGATGGGAGTGGGATCAGGTGGTTCAAATCTTATACACGGAGAATCCTAGG 277

 146 IleThrGlyAspGluLeuAsnGln 153
     ||||||:::|||||||||:::
 278 GTTTTAGATGATGAGAATTTAGATAA 301

  51 PheGlyLeuVal................GluSerSerThrProAsnAlaIle 62
     |||||| ||                   |||:|::||||||||||:||||||
   2 TTTGGTCAAGTAGCCTATTCTGCGGATGAGAGAGACTGTGCCGAATTTAG 51

  62 eAsnProAspSerSerSerGluTyrArgTrpTyrArgGlyTyrGluSerTyrV 79
     ::|||||||| |||:|||||||||||||||||||||||||:|||::|||
  52 TAGTCCA......AATCCTGAGTTCCCTTGGTATGGTTATGATGCTTAT. 94

  79 alArgGlyHisProTyrTyrLysGlnPheArgValAlaHisAspLeuArg 95
     |||                 |||||||||||||||||||||||||:::
  95 ..AGAGGT..........TCCTTTCTGAGATATCATGATTTAAAT 127
```

Left Column

```
     96 ValAsnLeuGluGlySerArgSerTyrGlnValTyrCysPheAsnLeuGly 112
        ||||||||||||||||||||||||||||||||||||||||||
    128 GTTAATTTGGAAGGTACGCACGTCTTACCAAGTTATTTATTCGTTAATTTAGT 177

    112 sLysAlaPheProLeuGlySerAspSerSerValLysLysTrpTyrTyrL 129
        |||               |||::|||::||||
    178 AAGACAGGAACCGAGTAAGGTTAATGGTTTAAGAAAAAATTGGTTCAAAA 227

    129 ysHisAspGlyIleSerThrLysPheGluAspTyrAlaMetSerProArg 145
        |||::||||||||||::|||::::||::|||::||
    228 AACGTTGATGGTGATAATGCTCGTGTTGTTTAAAAAAATATGCAGCTAATCCTAGA 277

    146 IleThrGlyAspGluLeuAsnGlnLysLeu 155
        :::::::::::::::::::::::::
    278 GTRAATTGATGGGAGATTTAGAACGTAACATT 307
```

seq_documentation_block:

```
seq_name: gb_bal:AF009915

LOCUS           AF009915         309 bp      DNA             BCT       13-FEB-1998
DEFINITION      Streptococcus pyogenes fibronectin binding protein F gene, partial
                cds.
ACCESSION       AF009915
VERSION         AF009915.1   GI:2267183
KEYWORDS        .
SOURCE          Streptococcus pyogenes.
  ORGANISM      Streptococcus pyogenes
                Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
                Streptococcus.
REFERENCE       1  (bases 1 to 309)
  AUTHORS       Katerov,V., Andreev,A., Schalen,C. and Totolian,A.A.
  TITLE         Protein F, a fibronectin-binding protein of Streptococcus pyogenes,
                also binds human fibrinogen: isolation of the protein and mapping
                of the binding region
  JOURNAL       Microbiology 144 (Pt 1), 119-126 (1998)
  MEDLINE       98129085
REFERENCE       2  (bases 1 to 309)
  AUTHORS       Katerov,V.E., Andreev,A.S., Schalen,C. and Totolian,A.A.
  TITLE         Direct Submission
  JOURNAL       Submitted (23-JUN-1997) Medical Microbiology, Lund University,
                Solvegatan 23, Lund 22362, Sweden
FEATURES             Location/Qualifiers
     source          1..309
                     /organism="Streptococcus pyogenes"
                     /strain="M12 strain 73195/2110"
                     /db_xref="taxon:1314"
     CDS             <1..>309
                     /codon_start=2
                     /transl_table=11
                     /product="fibronectin binding protein F"
                     /protein_id="AAC38162.1"
                     /db_xref="GI:2267184"
                     /translation="FGQVAYSADEKTVPNFKSPDPDYPWYGYSYRGIFARYHNLKVN
                     LKGSKEYQAYCFNLTKYFPRPTYSTTNNFYKKIDGSGSAFKSYTANPRVLDENLDKLE
                     "
BASE COUNT      110 a        46 c        58 g        95 t
ORIGIN
```

alignment_block:

```
US-09-494-297-2 x AF009915

Align seg 1/1  to: AF009915  from: 1  to: 309    ..

alignment_scores:
        Quality:   209.50          Length:   309
        Ratio:     2.831           Gaps:       4
Percent Similarity: 68.519    Percent Identity: 42.593

     51 PheGlyLeuVal.................GluSerSerThrProAsnAlaIl 62
        |||||||  |||          ||:::::||::||:||||
      2 TTTGGTCAAGTAGCCTATTCTCGCGGATGAGAGGAGAGACTGTCCGAAATTTAA 51
```

Right Column

```
     62 eAsnProAspSerSerGluTyrArgTyrIrgTyrIrgGlyTyrGluSerTyrV 79
        :::::|||||||  ::|||||| ||||||||||||||| |||||||
     52 AAGCCCAGAT..........CCAGATTATCCCTGGTATGGTTATTATTCGTTAT.  94

     79 aIArgGlyHisProTyrTyrLysGlnPheArgValAlaHisAspLeuArg 95
        |||                      |||:||||||::||||
     95 .AGAGGA..............ArATTTGCAAGATATCACAATTTAAA 127

     96 ValAsnLeuGluGlySerArgSerTyrGlnValTyrCysPheAsnLeuY 112
        |||                  |||||||:||:||||:||||
    128 GTAAATCTAAAAGGAAGTAAGGAGTATCAACGGTAATTGTTTAACCTAAC 177

    112 sLysAlaPheProLeuGlySerAspSerSerValLysLysTrpTyrTyrLysL 129
        :||               ||||::||||:|||||||
    178 AAAATACTTTCCTCGCCCCACTYATAGTACTACAAAATAATTTTTACAAGA 227

    129 ysHisAspGlyIleSerThrLysPheGluAspTyrAlaMetSerProArg 145
        ||           |||::::||:||||::|||||||
    228 AAATTGATGGGAGTGGATCAGCGTTCAAAATCTTATACAGCGGAATCCTAGG 277

    146 IleThrGlyAspGluLeuAsnGln 153
        |||::::|||||||
    278 GTTTTAGATGAGAATTTAGATAAA 301
```

seq_documentation_block:

```
seq_name: gb_bal:AF009912

LOCUS           AF009912         282 bp      DNA             BCT       13-FEB-1998
DEFINITION      Streptococcus pyogenes fibronectin binding protein F gene, partial
                cds.
ACCESSION       AF009912
VERSION         AF009912.1   GI:2267177
KEYWORDS        .
SOURCE          Streptococcus pyogenes.
  ORGANISM      Streptococcus pyogenes
                Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
                Streptococcus.
REFERENCE       1  (bases 1 to 282)
  AUTHORS       Katerov,V., Andreev,A., Schalen,C. and Totolian,A.A.
  TITLE         Protein F, a fibronectin-binding protein of Streptococcus pyogenes,
                also binds human fibrinogen: isolation of the protein and mapping
                of the binding region
  JOURNAL       Microbiology 144 (Pt 1), 119-126 (1998)
  MEDLINE       98129085
REFERENCE       2  (bases 1 to 282)
  AUTHORS       Katerov,V.E., Andreev,A.S., Schalen,C. and Totolian,A.A.
  TITLE         Direct Submission
  JOURNAL       Submitted (23-JUN-1997) Medical Microbiology, Lund University,
                Solvegatan 23, Lund 22362, Sweden.
FEATURES             Location/Qualifiers
     source          1..282
                     /organism="Streptococcus pyogenes"
                     /strain="T27 strain SF40"
                     /db_xref="taxon:1314"
     CDS             <1..>282
                     /codon_start=2
                     /transl_table=11
                     /product="fibronectin binding protein F"
                     /protein_id="AAC38159.1"
                     /db_xref="GI:2267178"
                     /translation="FGQVAYAADERTVPNHSSPNPEFPWYGYDSYRGIFARYHNLKVN
                     LKGSKEYQAYCFNLTKYFPRPTYSTTNNFYKKIDGSGSAFKSYAANPRV"
BASE COUNT       91 a        45 c        58 g        88 t
ORIGIN
```

alignment_block:

```
alignment_scores:
        Quality:   205.50          Length:   101
        Ratio:     2.978           Gaps:       4
Percent Similarity: 68.317    Percent Identity: 44.554
```

US-09-494-297-2 x AF009912

Align seq 1/1 to: AF009912 from: 1 to: 282

```
 51 PheGlyLeuVal...............GluSerSerThrProAsnAlaIl 62
    |||||||                    |||::::::::||||||||
  2 TTTGGTCAACTAGCCTATGCTGCGGATGAGAGACTGTGCCGAATCATAG 51

 62 eAsnProAspSerSerGluTyrArgTrpTyrGlyTyrGluSerTyrV 79
    ::|||   :    |||::: ||||||||||||||:||:::::::||
 52 TAGTCCA......AATCCTGAGTTCCTTGGTATGGTTATGATTCGTAT. 94

 79 aIArgGlyHisProTyrTyrLysGlnPheArgValAlaHisAspLeuArg 95
    |||||||   :::       |||:||||||||||||||||||:::
 95 .AGAGGA............ATATTTGCAAGATATCACAATTTAAAA 127

 96 ValAsnLeuGluGlySerArgSerTyrGlnValTyrCysPheAsnLeuY 112
    ||||||||:||||||||||||||||:||||||||||||||||||||
128 GTAAAACTAAAAGGAAGTAAGGAGTATCAAGCGTATTGTTTTTAACCTAAC 177

112 sLysAlaPheProLeuGlySerAspSerSerValLysLysTrpTyrLysL 129
    |||           |||        |||||:||||||||||:::
178 AAAATACTTTCCTCGCCCCACTATAGTACTACACAAAAAATTTTTACAAGA 227

129 ysHisAspGlyIleSerThrLysPheGluAspTyrAlaMetSerProArg 145
    |||||   ::::    :::|||||||  ::|||||
228 AAGTTGATGGTCATAATCCGTGTGTTTAAAAAATATGTAGCCTAATCCTAGA 277

146 Ile 146
    :::
278 GTT 280
```

seq_name: gb_bal:AF009919

FEATURES   Location/Qualifiers
     source   1..308
                     /organism="Streptococcus pyogenes"
                     /strain="M48 strainB4C3/48/1"
                     /db_xref="taxon:1314"
     CDS   <1..>308
                     /codon_start=2
                     /transl_table=11
                     /product="fibronectin binding protein F"
                     /protein_id="AAC38166.1"
                     /db_xref="GI:2267192"
                     /translation="FGQVAYAADEKTVPNYSSPNPEFPWIGYDAYRGWFLRYHDLNVN
                     LEGSASYQVCFNLVRQEPSKVNGLRKNWFKKVDGHNAVFKKYVANPRVIDGDLERNI
                     "

BASE COUNT   97 a   36 c   71 g   104 t
ORIGIN
```

US-09-494-297-2 x AF009919